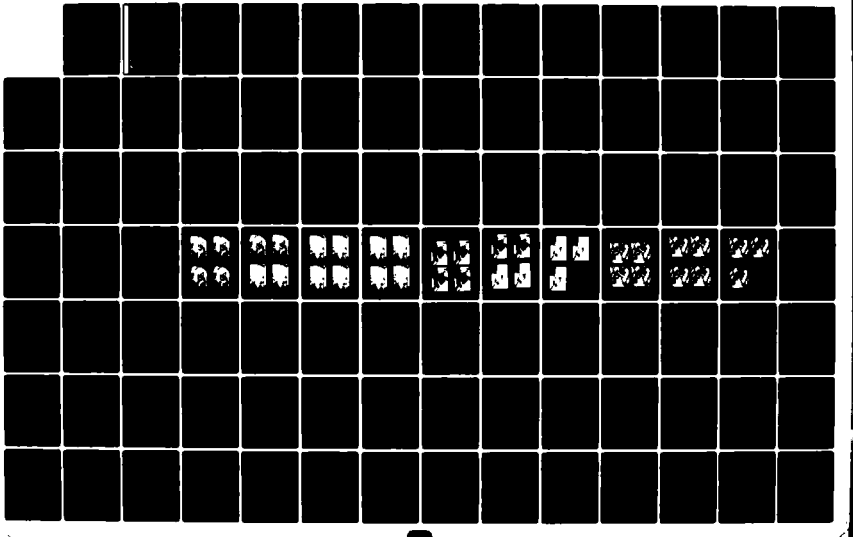
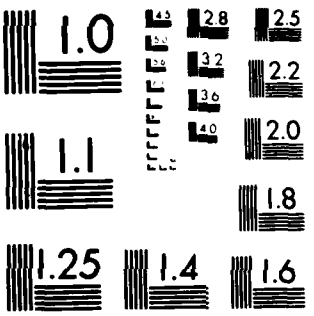
AD-A129 221 ADAPTIVE HYBRID PICTURE CODING(U) ARKANSAS UNIV 1/2
FAYETTEVILLE DEPT OF ELECTRICAL ENGINEERING
R A JONES ET AL. 05 FEB 83 AFOSR-TR-83-0499
UNCLASSIFIED AFOSR-77-3456 F/G 12/1 NL





UNIVERSITY OF ARKANSAS

ADAPTIVE HYBRID PICTURE CODING

AFOSR-TR- 83-0499

Final Report

by

Richard A. Jones
Principal Investigator

Carl D. Bowling
Senior Investigator

Yogendra Tejwani
Senior Investigator

Department of Electrical Engineering
University of Arkansas
Fayetteville, Arkansas 72701

February 5, 1983

Prepared for the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Under Grant No. AFOSR 77-3456

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 83-0499	2. GOVT ACCESSION NO. AD-A129 221	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ADAPTIVE HYBRID PICTURE CODING		5. TYPE OF REPORT & PERIOD COVERED FINAL REPORT 30 SEP 1977 to 01 OCT 82
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Richard A. Jones		8. CONTRACT OR GRANT NUMBER(s) AFOSR-77-3456
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Arkansas Department of Electrical Engineering Fayettevill, AK 72701		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2305/B3
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research Building #410 Bolling AFB Washington, DC 20332		12. REPORT DATE 5 February 1983 13. NUMBER OF PAGES 110
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release;
distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report consist of two parts. In part one, a time modified autoregressive model for interframe image coding is presented. This method is compared with previous work in the field of interframe image coding and it is shown that substantial simplifications occur when the nearest integer displacement is taken into account. It is demonstrated that when the between frame noise is minimal and the motion is pure translation or can be modelled by translation,

-OVER

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

or can be modelled by translation, enough information can be extracted from the predictor coefficients to determine the non-integer displacement with small error. In part two, a new concept for examining shapes as vectors in a shape space is described. The shape space is defined in terms of its properties and the importance of the independence of the size variable to the shape vectors, defined on this shape space, is stressed. Also, two theorems helpful in the process of comparing partial shapes to the complete shape are stated and proved. A new method for detecting the points on a shape which appear to dominate visual perception is described. This method, called the Adaptive Line of Sight Method detects the dominant points on a shape even though they do not always occur on points of high curvature. With this method, the critical points, or dominant points, of the shape that are determined are based on a set of coordinate axes that are dependent on the shape itself. Therefore, the points determined are independent of size, rotation, or relative displacement. The Line of Sight of a point concept is also introduced and subsequently utilized to extract features from a shape. These features are then compared to the features of other shapes by a syntactic procedure for the purpose of recognizing whether a slope is a partial shape or is a shape in its own right. It is demonstrated that the feature vectors determined by this producure are independent of size, rotation, and displacement. The results of applying these techniques to actual shapes are demonstrated and discussed.

FINAL REPORT

AFOSR-
GRANT NO. 77-3456

ADAPTIVE HYBRID PICTURE CODING

Richard A. Jones
Principal Investigator

Carl D. Bowling
Senior Investigator

Yogendra Tejwani
Senior Investigator

University of Arkansas
Department of Electrical Engineering
Fayetteville, Arkansas 72701

February 5, 1983

Prepared for the Air Force Office of Scientific Research

ABSTRACT

This report consist of two parts. In part one, a time modified autoregressive model for interframe image coding is presented. This method is compared with previous work in the field of interframe image coding and it is shown that substantial simplifications occur when the nearest integer displacement is taken into account. It is demonstrated that when the between frame noise is minimal and the motion is pure translation or can be modelled by translation, enough information can be extracted from the predictor coefficients to determine the non-integer displacement with small error.

In part two, a new concept for examining shapes as vectors in a shape space is described. The shape space is defined in terms of its properties and the importance of the independence of the size variable to the shape vectors, defined on this shape space, is stressed. Also, two theorems helpful in the process of comparing partial shapes to the complete shape are stated and proved. A new method for detecting the points on a shape which appear to dominate visual perception is described. This method, called the Adaptive Line of Sight Method detects the dominant points on a shape even though they do not always occur on points of high curvature. With this method, the critical points, or dominant points, of the shape that are determined are based on a set of coordinate axes that are dependent on the shape itself. Therefore, the points determined are independent of size, rotation, or relative displacement.

The Line of Sight of a point concept is also introduced and sub-

sequently utilized to extract features from a shape. These features are then compared to the features of other shapes by a syntactic procedure for the purpose of recognizing whether a slope is a partial shape or is a shape in its own right. It is demonstrated that the feature vectors determined by this procedure are independent of size, rotation, and displacement. The results of applying these techniques to actual shapes are demonstrated and discussed.

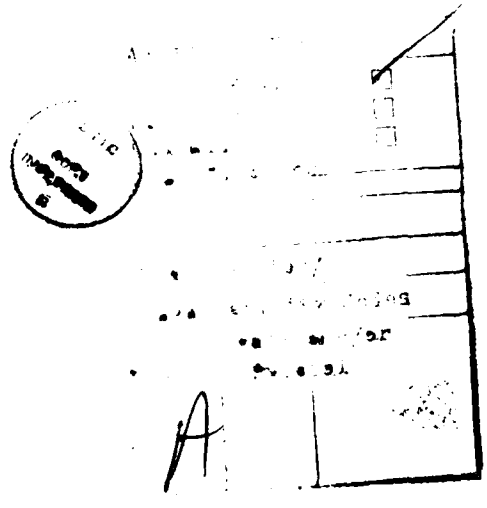


TABLE OF CONTENTS

LIST OF FIGURES

PART ONE

DISPLACEMENT ESTIMATION BY PREDICTION COEFFICIENT ENERGY CONCENTRATION

PROBLEM STATEMENT

Many scientific disciplines have sought solutions for questions associated with motion. What is motion? How is motion defined? How is motion interpreted? These are just a few of the endless number of questions that arise when dealing with the topic of motion. Psychologists look to the internal thought processes to formulate theories about human interpretation of motion through visual stimulation. Physiologists on the other hand are concerned with what biological processes are necessary for the generation of the synaptic signals associated with the sensations of vision and motion. Engineers then try to simulate the processes of biological vision in machines.

From research carried out in the above disciplines, it has been found that the human visual system is a very complex and complicated network of biological subunits. Some examples of these subunits are the light receptors, i.e. the rods and cones, the optic nerve for visual transmission, and the brain with associated memory for interpretation. Each of the subunits themselves constitute a very complex system. So it is no suprise to find that when the human visual system is modeled by hardware and software that the non-biological visual system will also be a very complex and complicated set of subunits.

The research effort detailed within deals with one very small subunit of the visual process, namely displacement or motion detection and estimation.

DISPLACEMENT AND MOTION

Simply stated, motion is defined to be a time series of spatial displacements. That is, in order for motion to be perceived, time must pass. If artificial vision and intelligence is to ever become a reality, then a sufficiently good mathematical model for motion will have to be employed. For this reason and the fact that memory space will always be limited, the vision system for motion should be based on some time adaptive displacement algorithm.

The problem then can be stated: Find a method to determine spatial displacement from an image sequence such that an estimate of the direction and magnitude of any detected motion can be made.

The problem statement is simple enough, but the effort is complicated by many factors. One of these factors involves object-background and object-foreground interaction. For example, if an object in the input frame moves in such a way so as to uncover some background, complications will arise in that the new information now consists not only of the object motion but also in the new background that is uncovered. The system must have the capacity to flag the

difference between the moving target and the non-moving background. Another problem simpler in scope than the above, but just as important, involves the loss of moving objects and the addition of new moving objects from outside the field of view. Again the system should be able to detect and track these new moving objects and discard the exiting objects.

Once the motion has been identified, there will be many uses for the information thus provided. Visual tracking of moving objects will then be possible as well as target trajectory prediction. The displacement vector need not be used strictly for motion related studies but may also be used in areas of data compression, remotely piloted vehicle control, and industrial manufacturing. With the detection and interpretation of motion, a very important step toward artificial vision will have been obtained.

Possible solutions to the problem that seem non-tractable at present may in the near future be made possible due to the advances in VLSI technology, software development, parallel processing and electo-optical systems. So even though the process may look overly complicated and slow at present there may yet be hope in the future.

PREVIOUS WORK IN THE FIELD OF MOTION

There has been relatively little work in the area of machine motion analysis until very recently when the required hardware and software became available. The work has concentrated in the areas of displacement estimation and interframe image coding. One of the earliest, and perhaps simplest, methods used for motion detection was simple image frame differencing. That is, subtract the previous frame, pixel by pixel, from the current frame and flag as motion any difference greater than some set threshold.

$$M(i,j,t) = ABS[I(i,j,t) - I(i,j,t-r)] \qquad (1)$$

Motion will be defined whenever M(i,j,t) is greater than some threshold. Although very simple, the method does show good results for a very limited class of simple images, but this method has many drawbacks that will limit its usefulness. First, the output is very sensitive to noise in the input images because it is a differentiating type process. Also any camera motion between image frames will translate into motion at every pixel. Finally, no information is available pertaining to the direction or magnitude of the motion.

Many other methods have been employed since the first frame differencing techniques. A method by Price, Snyder, and Rajala [74] uses a Fourier-Domain filter based on a

model of the human visual system to detect motion. The model divides visual perception into two distinct channels. The first, their so called x channel, is a temporal low pass and spatial band pass channel. This channel processes the information contained in two dimensional patterns with high spatial resolution but fairly low temporal dependance. They believe that it is this channel that is responsible for objects with structural complexity but with little or no motion. The other, so called y channel, is just the opposite. Its characteristics are spatial low pass and temporal bandpass. This channel, they claim, conveys the information of objects with high temporal dependance and low spatial resolution. It is this y channel that is used for the detection of motion.

Stuller, Netravali and Robbins of Bell Laboratories start from a completely different point of view for motion. First, the end product of their work is data compression and not tracking, although it could be modified for such. The model is used for normal televison data where adjacent scan rows are scanned by an interleaved method. In the first method of pel-recursive displacement estimation by Robbins and Netravali [78], the image model for pure translation with no background is given in equation 2.

$$I(X_k,t) = I(X_k - D, t - \tau) \qquad (2)$$

Where: $I(X_k,t)$ is the intensity of the image at the spatial location X_k and time t. $I(X_k - D,t - \tau)$ is the intensity of the image at the spatial location X_k adjusted by the displacement D at the previous time t- τ. They next define a displaced frame difference term as

$$DFD(X_k,D) = I(X_k,t) - I(X_k - \hat{D},t - \tau). \quad (3)$$

An attempt is made to minimize this difference with a steepest descent algorithm with the form of equation 4.

$$\hat{D}_{k+1} = \hat{D}_k - (1/2)\varepsilon\nabla\hat{D}_k[DFD(X_k,\hat{D}_k)]^2 \quad (4)$$

Where ε is a gain term and ∇_{D_k} is a two-dimensional gradient operator with respect to D_k Simplifying yields,

$$\hat{D}_{k+1} = \hat{D}_k - \epsilon\, DFD(X_k,\hat{D}_k)\,\nabla I(X_k - \hat{D}_k,t - \tau). \quad (5)$$

Following this same model, Netravali and Stuller [67] formulated a method for interframe coding termed coefficient recursive estimation. It is an extension of pel-recursive with the further addition of a unitary transform. The methods are similar, but now the image is broken up into rectangular blocks of size N_r rows by N_c columns. Each element is then multiplied by the appropriate transformation vector. The blocks are then changed into a column vector by column scanning the transformed block. They define the nth coefficient of the qth block of the transformed image to be,

$$c_n(q) = I^T(X_q,t)\phi_n \quad (6)$$

and for the estimated displaced frame,

$$c_n(q,\hat{D}) = I^T(X_q - \hat{D}, t - r)\phi_n \tag{7}$$

The comparable term for the previous method's displaced frame difference is the coefficient prediction error $e_n(q,D)$ and is given by equation 8.

$$e_n(q,\hat{D}) = [I(X_q,t) - I(X_q - \hat{D}, t - r)]^T\phi_n \tag{8}$$

The minimization is over the squared prediction error by a steepest descent iteration of the form in equation 9.

$$\hat{D}_{n+1}(q) = \hat{D}_n(q) - (\epsilon/2)\nabla_{\hat{D}_n(q)} e_n^2(q,\hat{D}_n(q)) \tag{9}$$

or in a simpler form,

$$\hat{D}_{n+1}(q) = \hat{D}_n(q) - \epsilon\, e_n(q,\hat{D}_n(q))G_n(q). \tag{10}$$

$G_n(q)$ is defined to be the coefficient gradient vector.

$$G_n(q) = [\nabla(I^T(X_q - \hat{D}_n(q), t - r)]\phi_n \tag{11}$$

When going to the next block the initial displacement estimate is set to the final estimate of the previous block.

$$\hat{D}_0(q) = \hat{D}_{N_r N_c - 1}(q-1) \tag{12}$$

When used for motion compensated interframe hybrid transform DPCM coding, the coder transmits a quantized version of the coefficient prediction error whenever it exceeds some threshold. This allows the reciever to update the estimate

of the displacement and also correct the prediction coefficients. Figures 1 and 2 point out the differences between a normal hybrid transform-DPCM coder-decoder pair, figure 1, and the motion compensated hybrid transform-DPCM coder-decoder pair, given in figure 2.

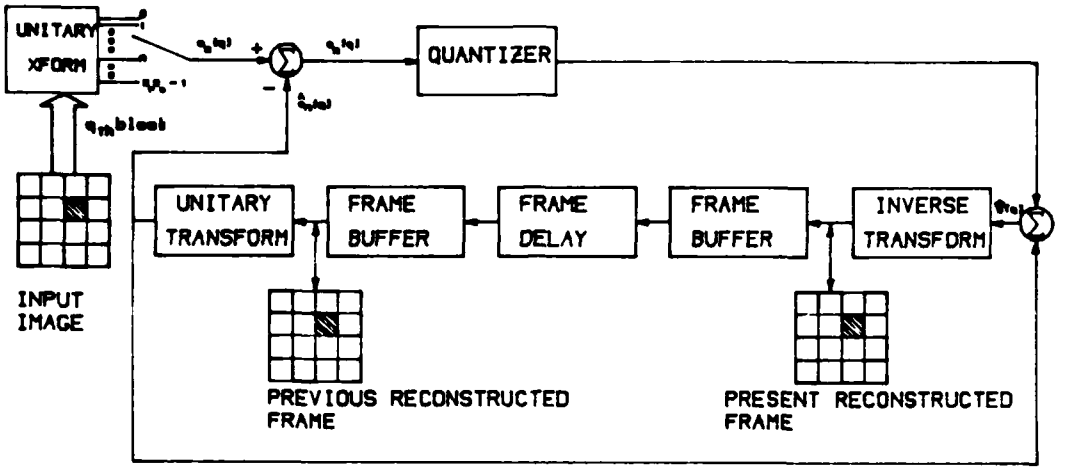


Figure 1. Normal Hybrid Transform-DPCM Coding

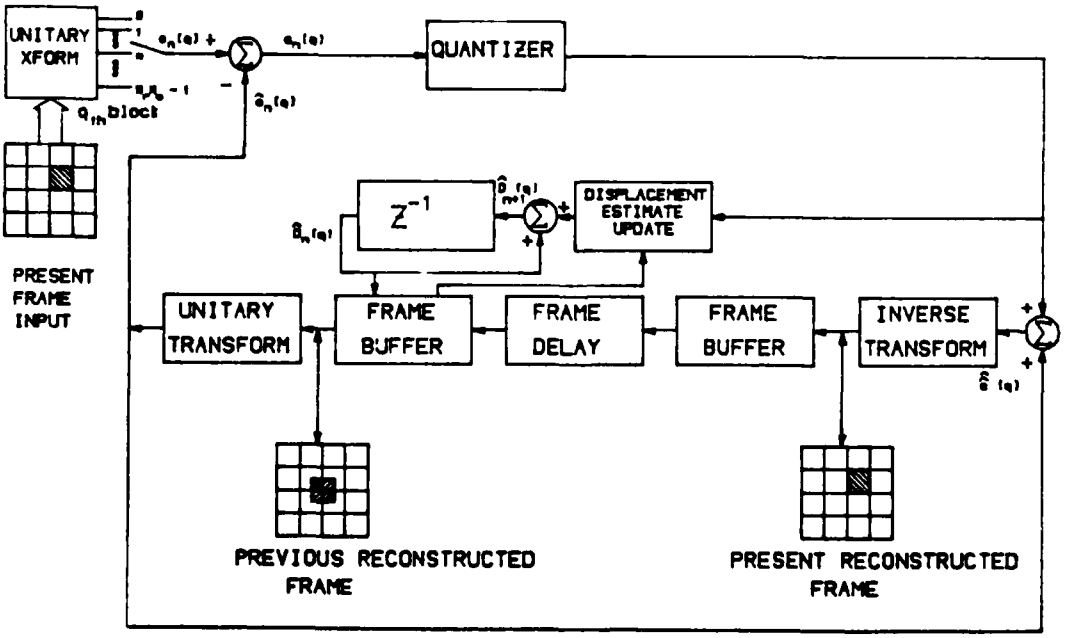


Figure 2. Motion Compensated Hybrid Transform DPCM Coding

Results of this and the earlier pel-recursive method are given in figures 4 through 11 for various values of gain. The test image was a radially decaying cosine function of radius 60 and peak-to-peak amplitude of 220 at the center and 130 at the circumference. The period also decreased radially starting with a period of 20 pixels at the center and ending with 10 pixels at the edge. The equation used to generate the test image is as follows.

$$I(R) = 100\exp(-0.01R)\cos(2\pi R/P) + 128 \qquad (13)$$

The displace image was displace in the x direction 2 pixel between time frames. Figure 3 is a picture of the actual data used.

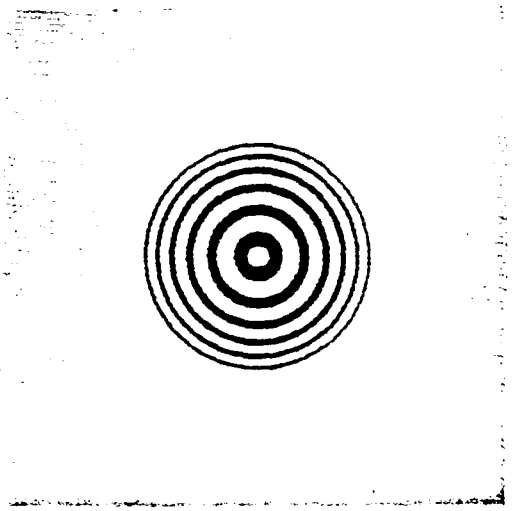


Figure 3. Radially decaying Cosine Function.

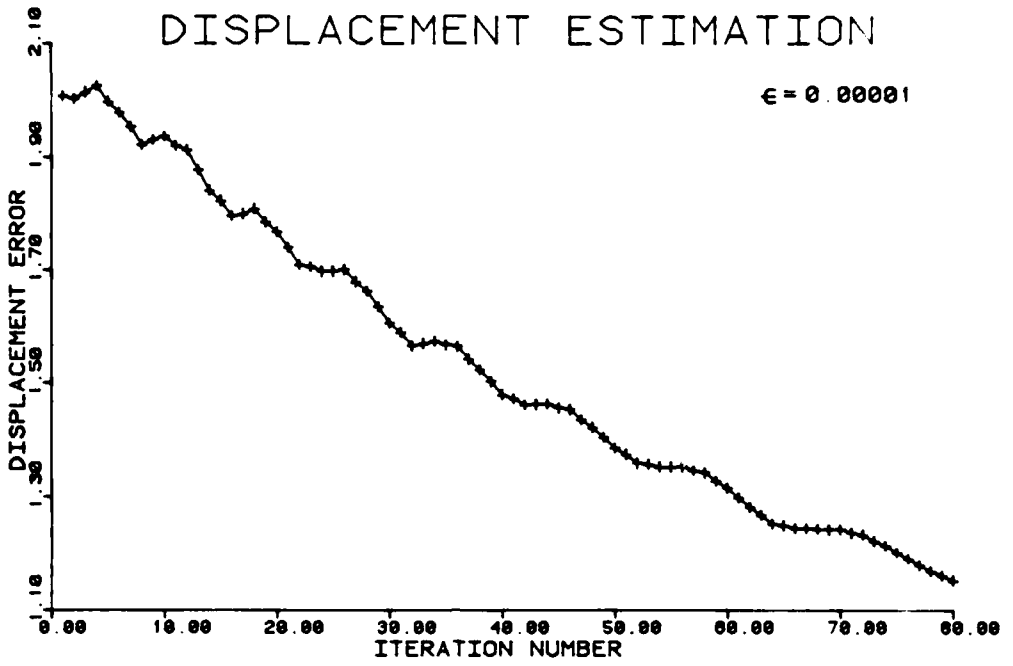


Figure 4. Pel Recursive Displacement Estimation.

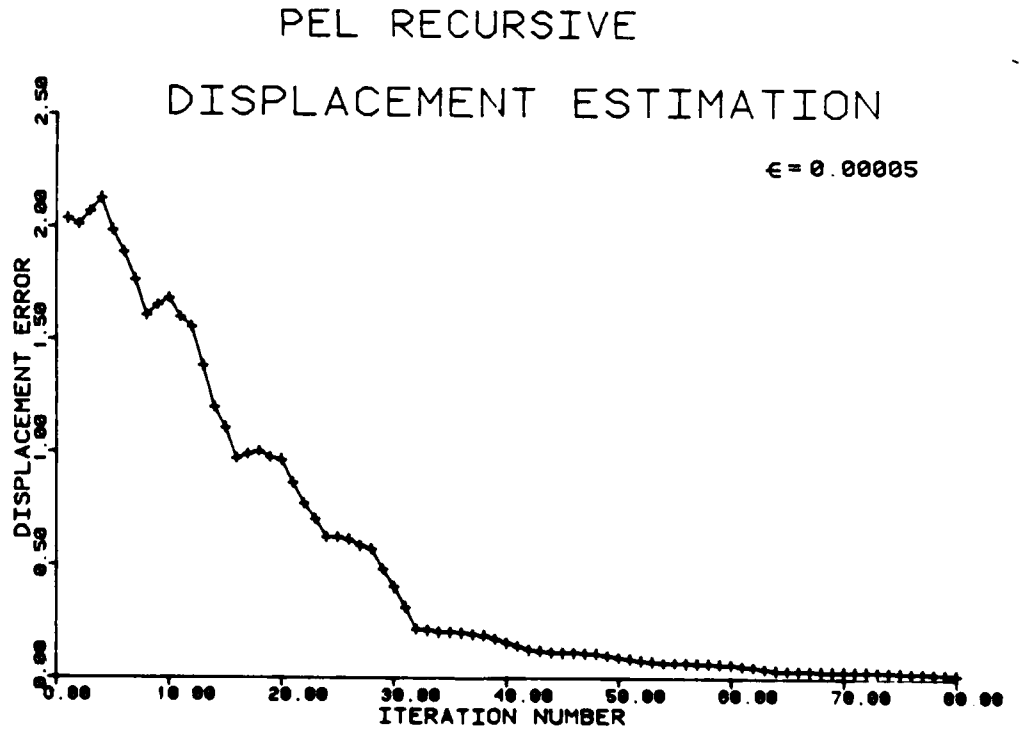


Figure 5. Pel Recursive Displacement Estimation.

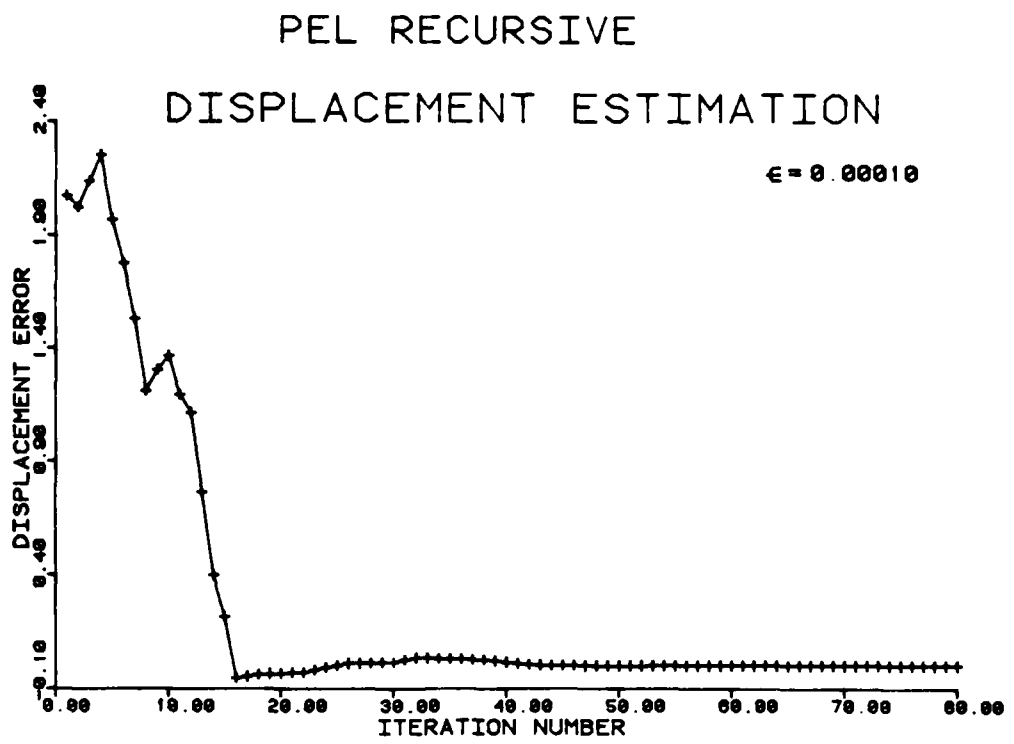


Figure 6. Pel Recursive Displacement Estimation.

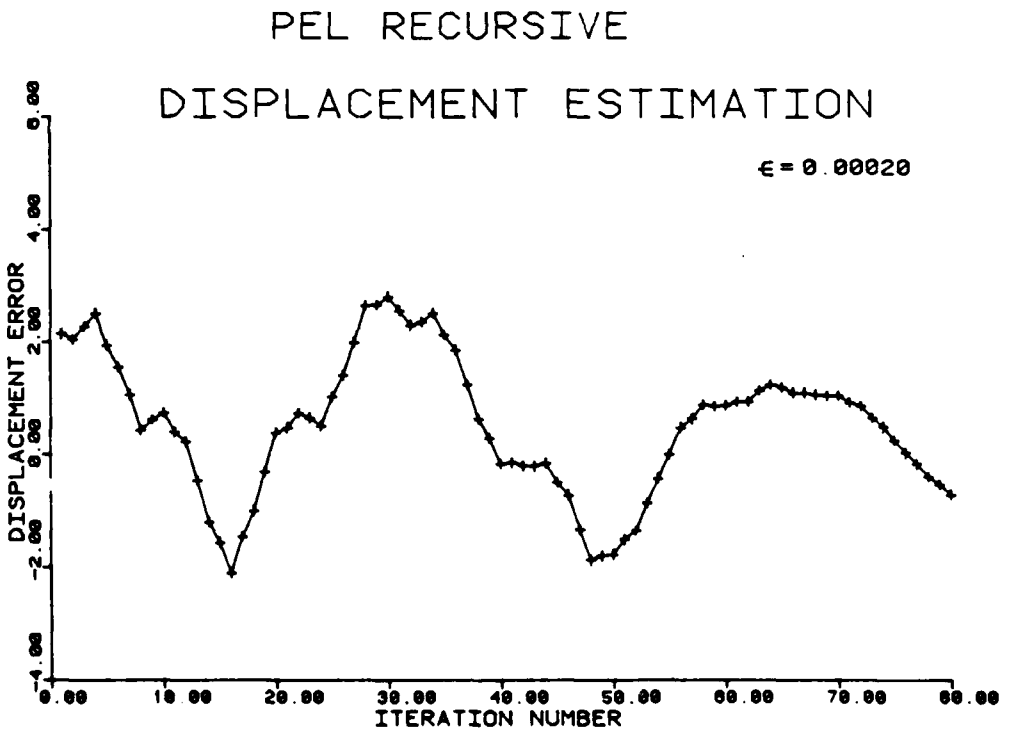


Figure 7. Pel Recursive Displacement Estimation.

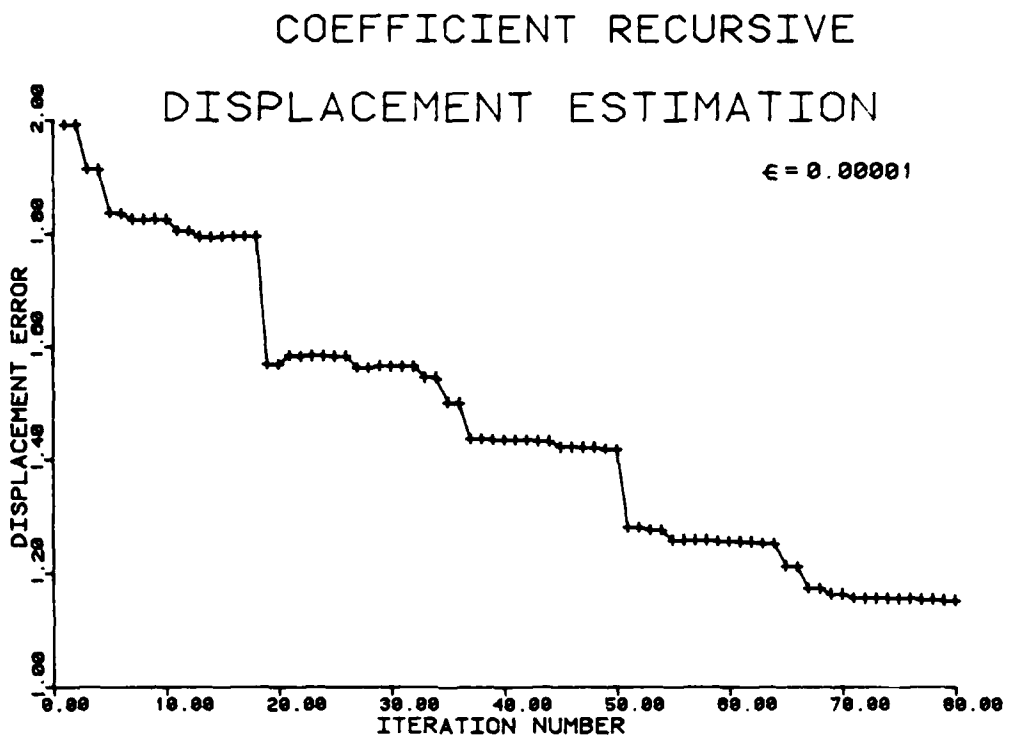


Figure 8. Coefficient Recursive Displacement Estimation.

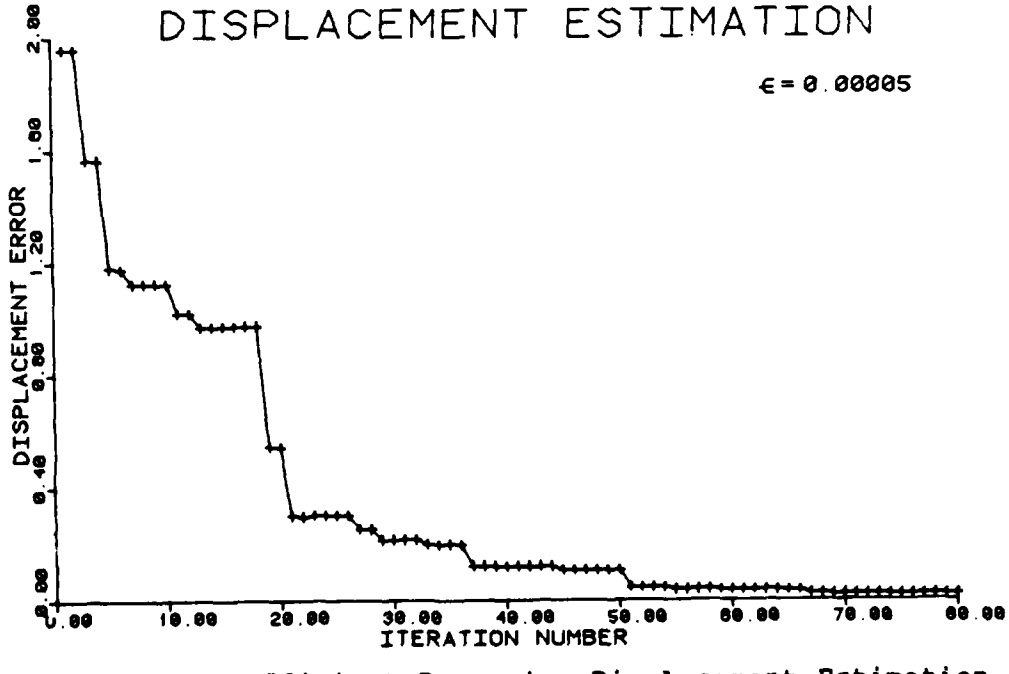


Figure 9. Coefficient Recursive Displacement Estimation.

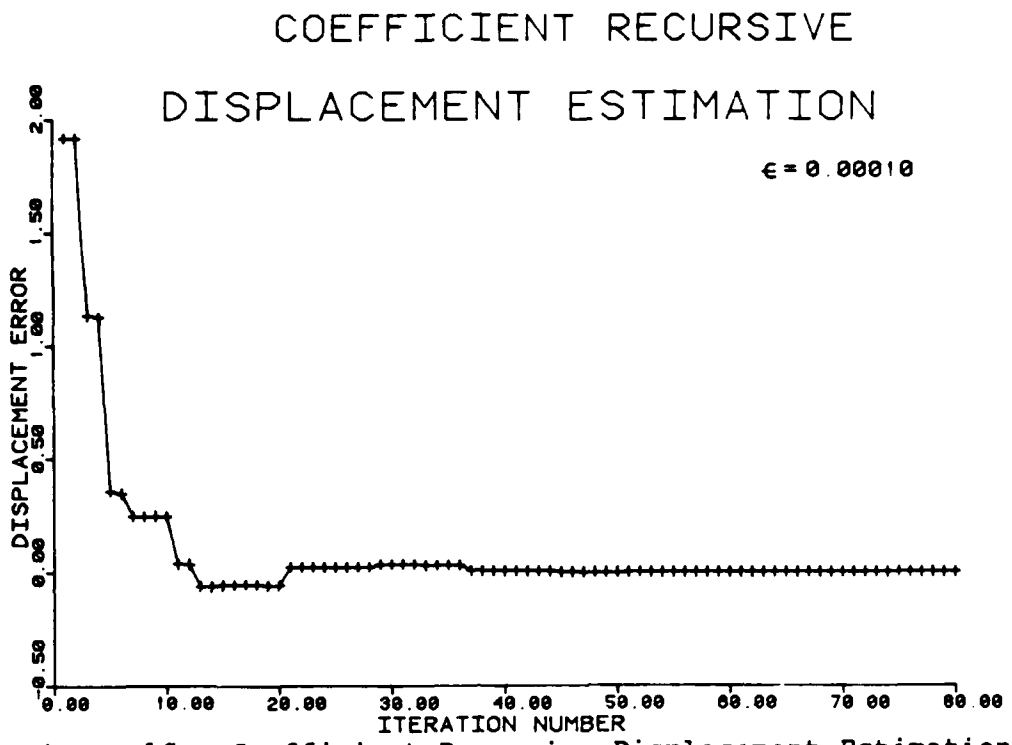


Figure 10. Coefficient Recursive Displacement Estimation.

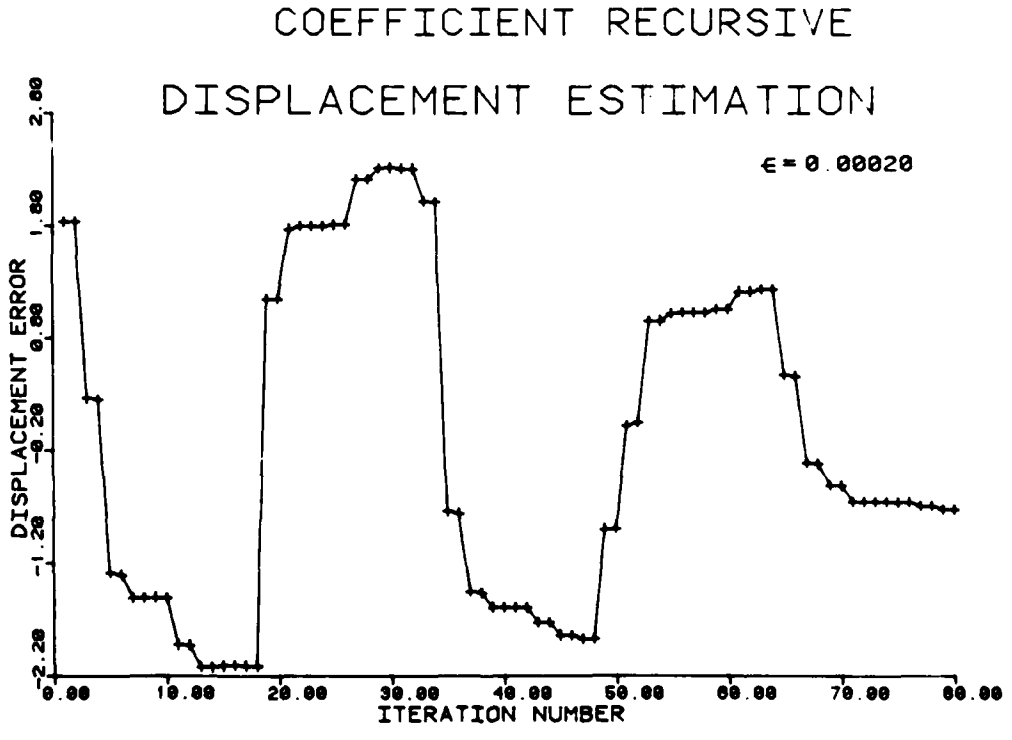


Figure 11. Coefficient Recursive Displacement Estimation.

The data compression that is achieved by the system can be attributed to two major characteristics. The first is due to the redundancy removal that is obtained by the prediction process and the second is due to the fact that many of the transform coefficients can be grossly quantized or completely neglected with acceptable results.

One of the major problems of coefficient recursive estimation is that it is very scene dependent. In one scene the displacement may converge in as few as 4 or 5 iterations where for others it may fail to converge at all. Another problem may prove to be the choice of a good value for the gain factor epsilon. As figures 4 through 11 verify, the choice of the gain factor plays an important role in the proper convergence of the algorithms.

PREDICTION COEFFICIENT ENERGY CONCENTRATION MODEL

Recalling again what the mathematical system is to perform and what the restrictions and fidelity criterion will be:

1. Be able to determine if motion has occured in the frame and if so where it has occured. Take the information provided about the motion and be able to track a target or recombine the needed data in such a way so as to form a good replica of the original input image.
2. Minimize the data required to update the system from one image frame to the next, that is achieve a fair amount of data compression while at the same time keeping the fidelity sufficiently high.
3. As a fidelity criterion or restriction, minimize the prediction error of both the displacement estimate and the reconstructed image estimate.

Working within these quidelines it can be seen tnat coefficient recursive displacement has met them all, but perhaps not to the extent that may be possible. The coefficient energy concentration model uses a similar motion model with some major modifications in the solution. The previous system exploited some of the available correlation of a two dimensional image, but by no means all. The correlation is not only available within the spatial domain, but also over the temporal domain. For maximum data

compression both spatial and temporal correlation should be exploited.

The image motion model is as follows. For the no background noiseless case define the current frame pixel intensity at the spatial location X_k to be a linear combination of pixels in the previous frame as shown in figure 12. The double crosshatched area represents a pixel in the current frame and the single crosshatched area represents the summation neighborhood in the previous frame. The summation neighborhood is the area in the previous frame whose pixel values are to be used for the prediction of the current frame pixel value.

$$I(X_k,t) = \sum_{Y_n} a(Y_n) I(X_k, t - r) \qquad (14)$$

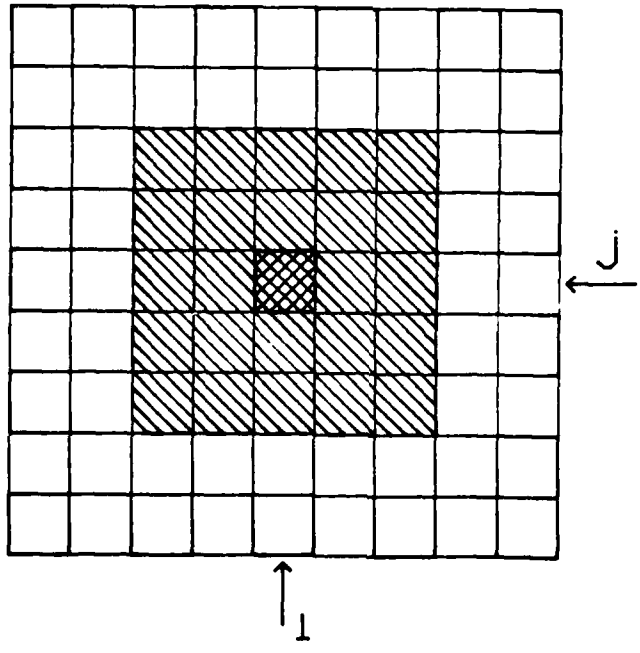


Figure 12. Two Dimensional Previous Pixel Neighborhood.

Rewriting X_k into the two spatial coordinates x and y yields,

$$I(x,y,t) = \sum_{-I}^{I}\sum_{-J}^{J} a(i,j)I(x - i,y - j,t - r). \qquad (15)$$

For two dimension non-separable motion this would require a matrix of (2I+1) by (2J+1) prediction coefficients for each pixel and hence little if any data compression could be obtained.

As stated earlier, it would be advantageous to capitalize on both the spatial and temporal correlation that exists in most image data. In this method the spatial correlation is exploited as a result of a linear combination being taken within a given time frame, and the temporal correlation is taken advantage of when the prediction is made from the past frame to the present.

Looking at the usual 2 dimensional autoregressive model, the previously defined model can be seen to be similar in many respects. The form is the same with the deletion of the prediction over time frames and the previos frame pixel neighborhood is different.

$$I(i,j) = \sum\sum a(k,l)I(i-k,j-l) + w(i,j) \qquad (16)$$

The values for i and j are allowed to vary over the picture size, the values for k and l are then able vary to take into account values above the current pixel and to the left in

the same line. Figure 13 shows the pixel neighborhood that may be used as in relation to that in figure 12.

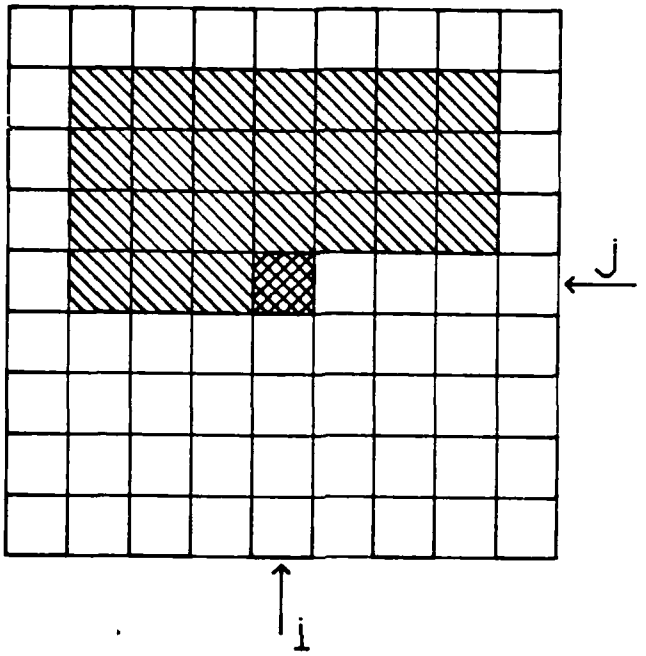


Figure 13. Current Frame Pixel Neighborhood.

In equation 16 the a(k,l) matrix contains the regression coefficients and w(i,j) then contains the error terms. Note that this is only able to take advantage of pixel values on one side of the current point, that is the values above it, as is shown in figure 13. Clearly this does not exploit the data to the fullest. Recalling that the model is normally used with a scanning system where the values to the right and below are not yet know by the reciever. Assuming full knowledge of the image at the previous time frame, a prediction over the time boundary yield the following form.

$$I(i,j,t) = \sum_{-K}^{K}\sum_{-L}^{L} a(k,l)\, I(i-k,j-l,t-r) + w(i,j) \qquad (17)$$

This in the sequel be termed the time modified 2 dimensional autoregressive model.

Noting what the physical implications are in relations to the regression problem, it can be seen that the coefficients are used to perform a translation and hence the model can be simplified. Assuming the image fields can be sampled infinitely fast such that they now become continuous functions, the regression coefficient matrix will translate to a delta function such that the above equation can be written as,

$$I(x,y,t) = \delta(\Delta_X,\Delta_Y) * I(x,y,t-\tau) + E(x,y) \qquad (18)$$

where the function * defines the convolution operator. Now the problem is to determine the values for Δ_X and Δ_Y or where the delta function is located. Here as in the previous methods, the theory works exactly only for the pure translation no background case, but approximations can be made and steps taken to improve its usefulness for non-ideal situations.

DETERMINING MAXIMUM ENERGY CONCENTRATION

Noting that if no noise exists then

$$I(x,y,t) = \delta(\Delta_x,\Delta_y) * I(x,y,t-\tau) \qquad (19)$$

should hold. Remembering that the sampled version I(i,j,t) is available, the above equation in discrete form becomes,

$$I(i,j,t) = \delta(\Delta_X, \Delta_Y) * I(i,j,t-\tau). \qquad (20)$$

Where the $\delta(\Delta_X, \Delta_Y)$ may not fall on the sampling points and hence amounts to convolving a continuous function with a discrete function. The countinuous function is simply a delta function at some displacement. The displacement of the delta function from the origin determines the image displacement from frame to frame. The problem is to solve for the location of this delta function. The method used to find the location of the delta function involves a two step process. First the integer part of the displacement is found via the use of a similarity metric and the non-integer part by a four coefficient prediction process.

THE SIMILARITY METRIC AND THE NON-INTEGER PORTION OF THE DISPLACEMENT

The similarity metric is used to find the best fit of the current frame image block with the previous image frame. This generates the nearest integer pixel displacement for the current block. The similarity metric that is used is based upon a non-linear combination of the current frame block and the previous frame block. A few different combinations have been tried with no real advantages or disadvantages exhibited by any one. One involves keeping

track of the sum of the absolute values of the differences for various values of shift.

$$M(i,j) = \sum_{-p}^{p}\sum_{-p}^{p} ABS[I(m-i,n-j,t-\tau) - I(m,n,t)] \qquad (21)$$

The values for i and j are set so as to vary over what is assumed to be the maximum displacement from frame to frame. The estimate for the integer displacement is defined to be the point which minimizes the metric M(i,j). One of the other possible metrics involves the sums of the squares of the differences for various shifts.

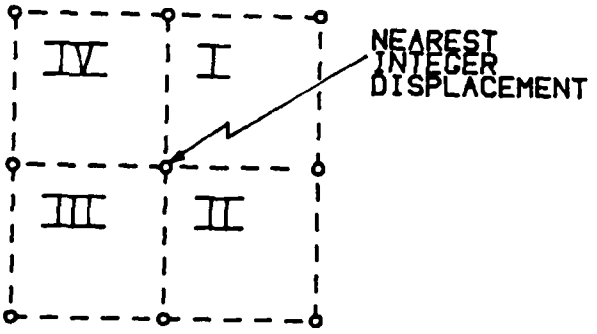
$$M(i,j) = \sum_{-p}^{p}\sum_{-p}^{p} [I(m-i,n-j,t-\tau) - I(m,n,t)]^2 \qquad (22)$$

Equation 22 offers the advantage of penalizing the metric greater for larger errors than for a number of smaller errors. Here again the minimum of the metric defins the estimate of the integer displacement. The metric that has proved to be the most benificial to this point is based on a counting procedure. As in the two above equations either the absolute value or square of the difference can be used. Further, a two value limiting function is define as L(x,y). Where L(x,y) is define to be 1 if x<y and 0 if x>y. This combined with the two equations yields a counting arrangement as follows.

$$M(i,j) = \sum_{-p}^{p}\sum_{-p}^{p} L(x, \quad ABS[I(m-i,n-j,t-\tau) - I(m,n,t)]) \qquad (23)$$

At this point the value for x is fairly arbitrary, but somewhat loosely related to the variance of the data noise and may even be a variable itself based on the data. Again as before, the minimum value of the metric M(i,j) will give the estimate of the integer part of the displacement.

Once an estimate has been determined for the integer portion of the displacement the problem then remains to determine where the delta function is in relation to the integer displacement. Figure 14 shows the 4 possible quadrants about the obtained integer displacement. For the noiseless case where only pure translation of the image field occurs, a similarity value indicating a perfect match will exist in one of the 4 quadrants.

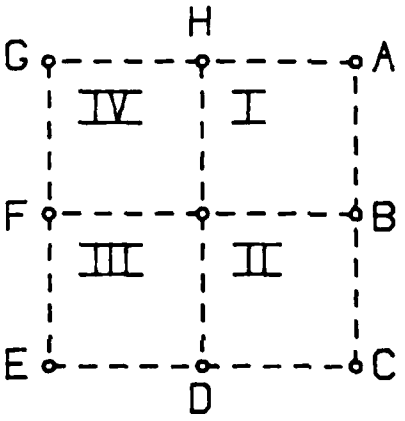


FOUR POSSIBLE QUADRANTS
ABOUT THE NEAREST INTEGER
DISPLACEMENT

Figure 14.

The determination of which quadrant is based on the values generated for the similarity matrix. Given that the similarity metric is a difference operator of some sort, the minimum indicates the quadrant containing the delta function.

The Minimum quadrant can be found by summing the values of the metric defining each quadrant as is shown in figure 15.



Q(1)=H+A+B
Q(2)=B+C+D
Q(3)=D+E+F
Q(4)=F+G+H

Figure 15.
LOCATION FOR NON-INTEGER PORTION OF DISPLACEMENT

Taking the minimum of Q(1) - Q(4) indicates the correct quadrant. This only narrows down the possibilities for the non-integer part, the actual calculations still remain. This is solved by assigning a coefficient to each corner of the indicated quadrant and solving for its value in a regression analysis. Each coefficient is multiplied by the corresponding previous image pixel value, the four are summed and set equal to the current image value as is indicated in equation 24.

$$I(m,n,t) = \sum_{i=0}^{1}\sum_{j=0}^{1} a(i,j)\, I(m-dx-i,n-dy-j,t-r) + a_{\emptyset} \qquad (24)$$

Where dx and dy indicate the integer X and Y displacement estimates respectively. Note that this is very similar to what was defined to be the two dimensional time modified autoregressive model. The only modifications are the summations have been narrowed to a single cell and the addition of the estimated integer displacement values dx and dy.

The problem then remains to find values for the predictor coefficients or a matrix. The solution is arrived at through the use of the normal regression approach. The problem can be further specified by considering it a two dimensional autoregressive model. The usual linear first order regression problem can be written.

$$Y_i = B_{\emptyset} + B_1 X_i + e_i \qquad (25)$$

This then tries to fit a straight line through the data for prediction purposes. The parameters of the model are $B_{\emptyset}$ and B_1 ,the independent variable is X and the dependent variable is Y. Values for $B_{\emptyset}$, B_1, and e are unknown and hence the problem is to determine estimates for these values. Given that $b_{\emptyset}$ is the estimate for $B_{\emptyset}$ and b_1 likewise for B_1 then the predictor equation can be written as follows.

$$Y = b_{\emptyset} + b_1 X \qquad (26)$$

To find the estimates for $b_\emptyset$ and b_1 the method of least squares is used, that is minimize the sum of the squares of the error term with respect to $B_\emptyset$ and B_1.

$$SSE = (Y_i - B_\emptyset - B_1 X_i) \tag{27}$$

Find the minimum of SSE with respect to $B_\emptyset$ and B_1. This is accomplished by taking the partials and setting equal to zero.

$$\frac{\partial SSE}{\partial B_\emptyset} = -2\sum_{i=1}^{n}(Y_i - B_\emptyset - B_1 X_i) = \emptyset \tag{28}$$

and

$$\frac{\partial SSE}{\partial B_1} = -2\sum_{i=1}^{n} X_i (Y_i - B_\emptyset - B_i X_i) = \emptyset \tag{29}$$

From these the estimate b_1 for B_1 is given by,

$$b_i = \frac{\sum_{i=1}^{N}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{N}(X_i - \overline{X})^2} \tag{30}$$

and the estimate $b_\emptyset$ for $B_\emptyset$ can be obtained from the following.

$$b_\emptyset = \overline{Y} - b_1 \overline{X} \tag{31}$$

With this another way to write the prediction equation is evident.

$$Y_i = \overline{Y} + b_1 (X_i - \overline{X}) \tag{32}$$

To make the notation simpler and somewhat easier to understand for larger systems it is often advantageous to

write the equations in matrix notation. The corresponding matrix equation for the above analysis is given by equation 33.

$$Y = XB + e \tag{33}$$

Y is the dependent variable vector of size n by 1. X is the augmented independent variable matrix of size n by 2. B is the parameter vector of size 1 by 2 containing the scaler terms $B_\emptyset$ and B_1. The matrix normal equation can be written as follows.

$$X^TXb = X^TY \tag{34}$$

The least squared estimate for b is then obtained from the following relation.

$$b = (X^TX)^{-1}X^TY \tag{35}$$

As stated earlier the model desired here is an autoregressive source. The autoregressive source places a further restriction on the regressive model in that the prediction is made with respect to a time series. That is future values of a time series are to be predicted on information provided by past values. The usual autoregressive source can be written as

$$X_{i+1} = B_\emptyset + B_1X_i + e_{i+1} \tag{36}$$

or in matrix notation,

$$X = ZB + e \tag{37}$$

Where Z represents a shifted version of X. In the particular case at hand, the derivation has to be carried one step further, it also has to take into account that the data is two dimensional. The two dimensional autoregressive model as stated earlier can be written as,

$$I(m,n,t) = \sum_{i=0}^{1}\sum_{j=0}^{1} a(i,j)\,I(m-dx-i,n-dy-j,t-r) + a_{\emptyset} + e(i,j) \qquad (38)$$

Notice that this is not in the ordinary matrix form. Some manipulation of the data is required in order to get the system into a form that is readily solvable by ordinary regression analysis. The method for transforming the two dimensional system into a single dimensional regression problem is as follows.

1. Rewrite the regression coefficient matrix a(i,j) into a regression vector B of length 5 where the first value comes from the intercept coefficient $a_{\emptyset}$.

$$B = \begin{bmatrix} a_{\emptyset} \\ a(\emptyset,\emptyset) \\ a(\emptyset,1) \\ a(1,\emptyset) \\ a(1,1) \end{bmatrix} \qquad (39)$$

2. Set up the independent variable matrix Z in such a way so as to take into account the shifting that is accomplished through the use of the double summation.

In the equation for the Z matrix that follows, the t-factor has been neglected for simplicity sake.

$$Z = \begin{bmatrix} 1 & I(i,j) & I(i,j+1) & I(i+1,j) & I(i+1,j+1) \\ 1 & I(i,j+1) & I(i,j+2) & I(i+1,j+1) & I(i+1,j+2) \\ & \cdot & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot & \cdot \\ 1 & I(i,j+J) & I(i,j+J+1) & I(i+1,j+J) & I(i+1,j+J+1) \\ 1 & I(i+1,j) & I(i+1,j+1) & I(i+2,j) & I(i+2,j+1) \\ & \cdot & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot & \cdot \\ 1 & I(i+I,j+J) & I(i+I,j+J+1) & I(i+I+1,j+J) & I(i+I+1,j+J+1) \end{bmatrix} \quad (40)$$

Figure 16 shows how the scanning of the previous data matrix is performed. Note that each value is used up to four times in the Z matrix.

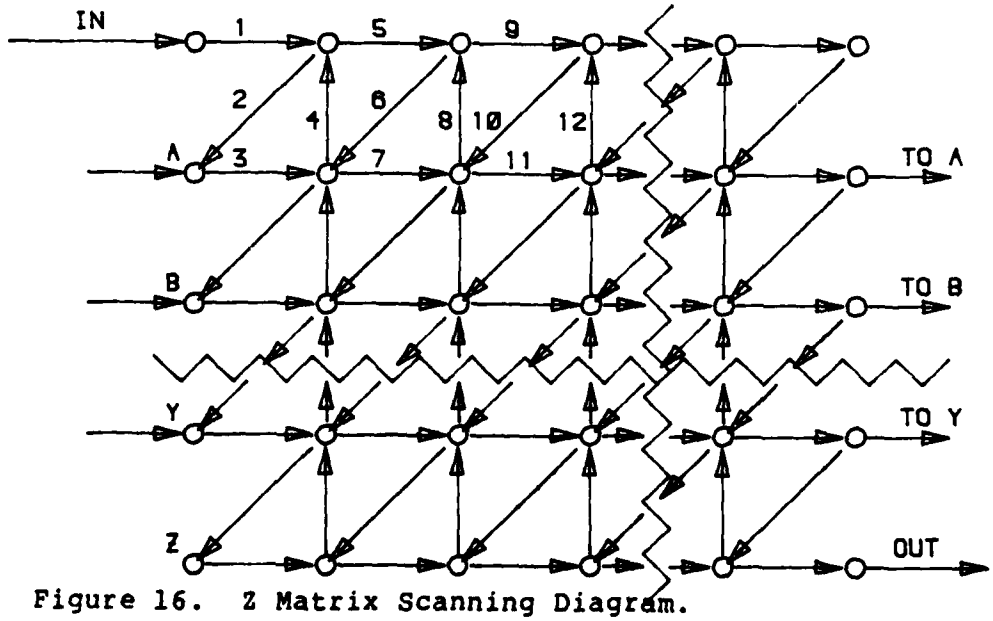


Figure 16. Z Matrix Scanning Diagram.

3. Finally row scan the current frame block data placing it in a column vector X of length I times J, where I and J are the x and y blocksize measurements of the comparison block. The scanning method is graphically shown in figure 17. Note also that the e vector would be defined identically to that of the X vector.

$$X = \begin{bmatrix} I(i,j,t) \\ I(i,j+1,t) \\ I(i,j+2,t) \\ \vdots \\ I(i,J,t) \\ I(i+1,j,t) \\ \vdots \\ I(i+I,j+J,t) \end{bmatrix} \qquad (41)$$

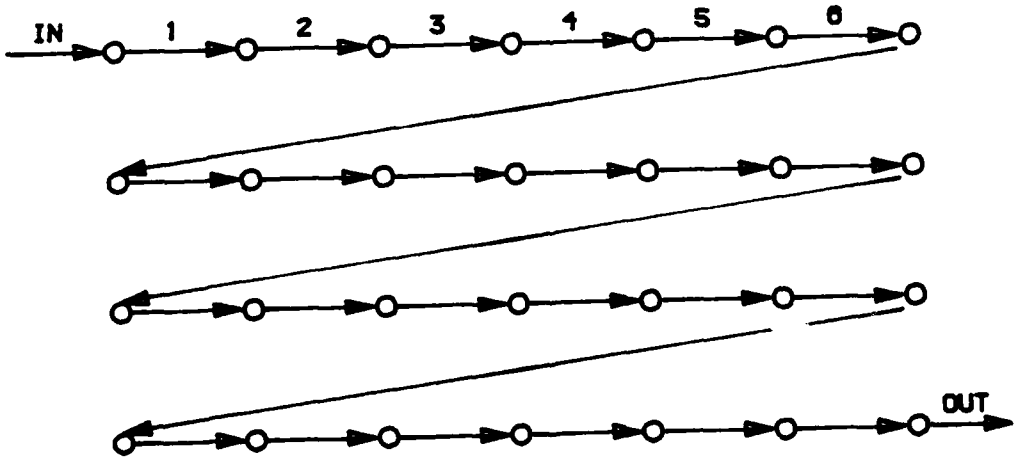


Figure 17. X Vector Scanning Diagram.

With each of the variables X,Z,B, ane e defined the current problem reverts to that of a first order regression problem involving five regression coefficients and hence can be solved as such. The predictor coefficients can hence be obtained by solving the following matrix equation for B.

$$X = ZB + e \tag{42}$$

The normal equation is then,

$$Z^T ZB = Z^T X \tag{43}$$

and the least squares estimate for B is b and is given by

$$b = (Z^T Z)^{-1} Z^T X. \tag{44}$$

As with any system that is based on obtaining a matrix inverse it is possible that the system is ill conditioned and the inverse may not exist. In the current system this is remidied by using only the estimate for the integer portion of the displacement if the current status proves the system to be ill conditioned. When this occurs the b vector is set to an identity transfer function, that is,

$$b^T = [0,1,0,0,0] \tag{45}$$

Solution of the regression problem will generate the four prediction coefficients and the intercept. For the pure

translation-noiseless case there is a very nice relationship amoung these coefficients. First, since the power in the past and present frames are nearly equal the coefficients will sum very close to one. Next since the coefficients act as a shifting mechanism each will be greater than zero. Finally because of the symmetry involved in the values of the four coefficient they can be thought of as being co-planer and the off sample point location of the delta function can be determined by finding the center of mass of the coefficients. Hence only the location of the delta function needs to be transmitted and the coefficients can be generated from this knowledge. When the motion is not pure translation and/or noise and/or non-zero background is present, then it advantageous to transmitt the values of the coefficients instead of the location of the delta function. With noise, rotation, scale change and/or other non-optimal occurances the nice features listed above may no longer hold. This is the reason for the transmission of the coefficient values themselves and not the delta function location. Hence for the data compression scheme, the quantized prediction coefficients and quantized prediction errors are transmitted. If the displacement is constant, then the coefficients need be transmitted only for the blocks where non-uniform displacement has occured. A block diagram if the system is given in figure 18 to show the order in which each of the steps is performed.

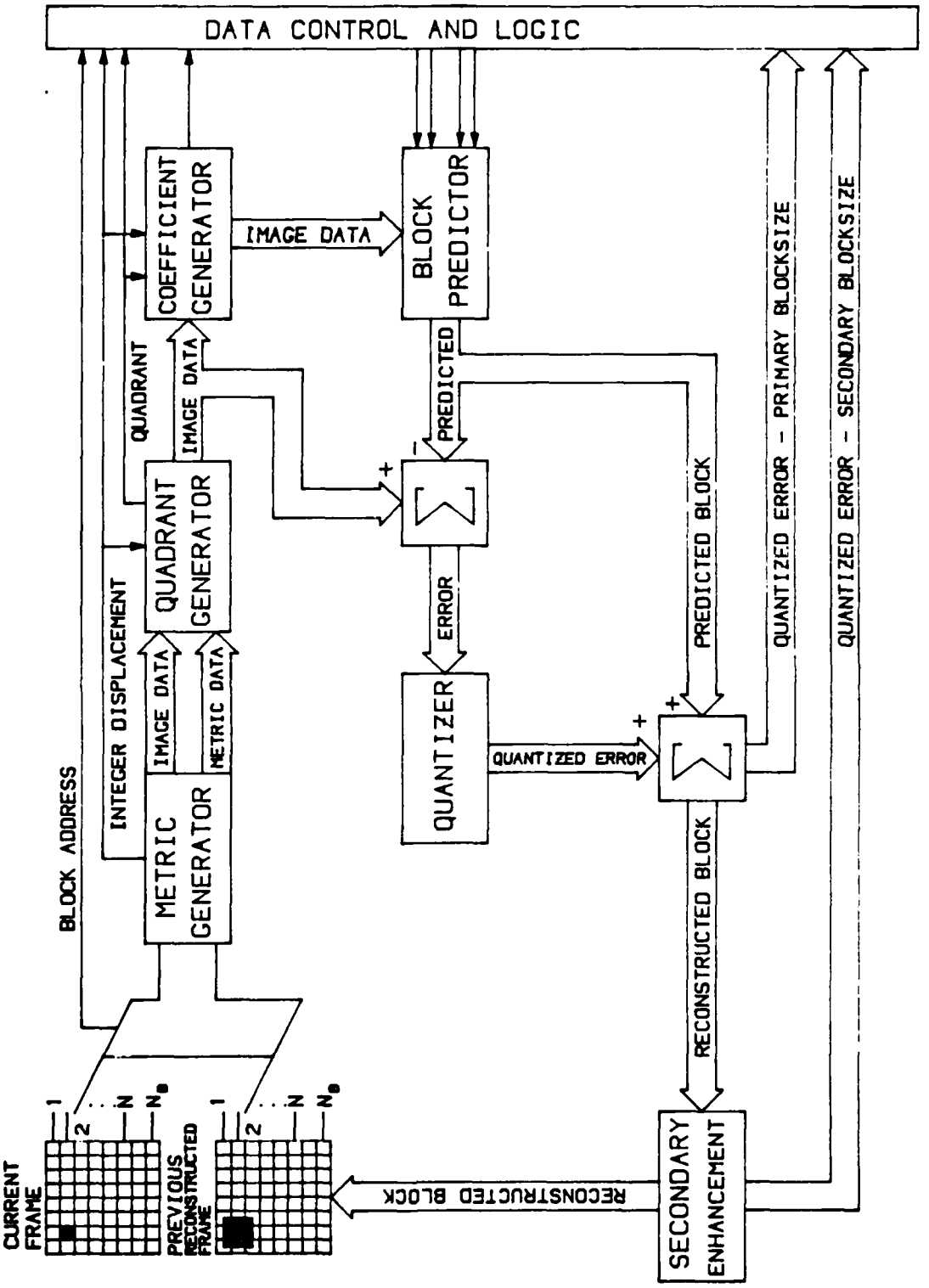


Figure 18. System Block Diagram

CONCLUSIONS AND RESULTS

With the assignment of a time modified autoregressive model for interframe image coding it has been shown that major simplifications can be made by taking into account nearest integer displacement. The limits on the summations for the regression no longer need be -p to p to resolve for a displacement of p pixel shift in either direction, but instead can be limited over a single cell. It has also been shown that for the translation only-noiseless case the prediction coefficients supply enough information to determine very closely the actual non-integer displacement. When combined with an appropriate coding scheme the method should produce respectable results with fewer computations than some of the methods previously proposed. It offers the advantages of not having to go through an iterative procedure to determine the displacement as well as being well suited for parallel implimentation. Noise in the images greatly affects the quality of the similarity metric and hence generates false maximum energy concentration points. The problems of rotation, scale changes, perspective changes and other non-translation types of motion can only be resolved with the addition of more intelligence to the system.

Three sets of test images sequences were used for the algorithm analysis. The first is a sequence of frames from

a digital television transmission. This particular sequence contains every other frame in the sequence so the problem of interlacing that the original data possessed could be neglected. The sequence consists of a woman talking on the telephone while at the same time making some rather quick motions with her head and hands. This particular sequence will be referred to as the 'FAST PHONE' in the remaining part of the discussion. The second image sequence was generated identically to that of the 'FAST PHONE' sequence but consisted of somewhat slower motion in the sequence. The final sequence consisted of an in-house motion sequence generation. A fixed background was placed under the vidicon and the motion was achieved by moving the small model plane some small distance between the image frame. These last two sequences will from here on be referred to as the 'SLOW PHONE' and'PLANE'.

Figures 19 a through f are the original 'FAST PHONE' frames while figures 20 a through e are the outputs with the associated bit rate required to transmitt.

Figure 19. "FAST PHONE" Input Sequence.

(a) Input Image 1.

(b) Input Image 2.

(c) Input Image 3.

(d) Input Image 4.

Figure 19 Continued.

(e) Input Image 5.

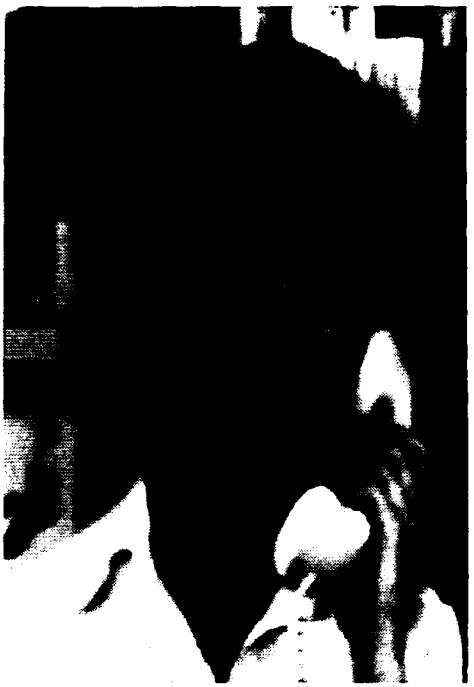
(f) Input Image 6.

Figure 20. Output Image Sequences.

(a) Output Image 2 (.154)

(b) Output Image 3 (.368)

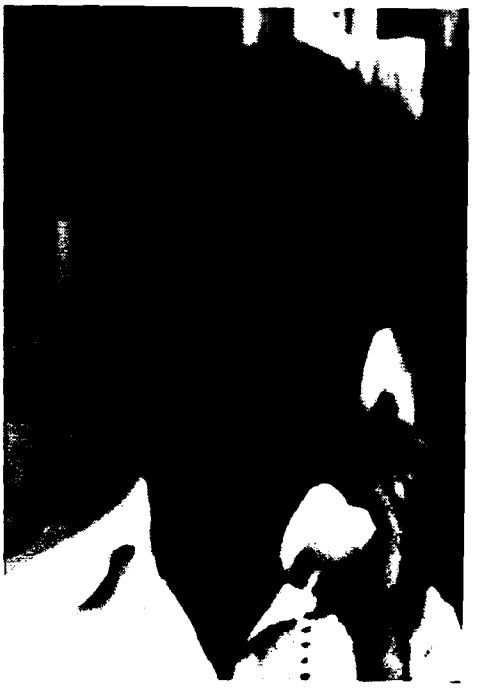
Figure 20 Continued.



(c) Output Image 4 (.584)

(d) Output Image 5 (1.00)

(e) Output Image 6 (1.42)

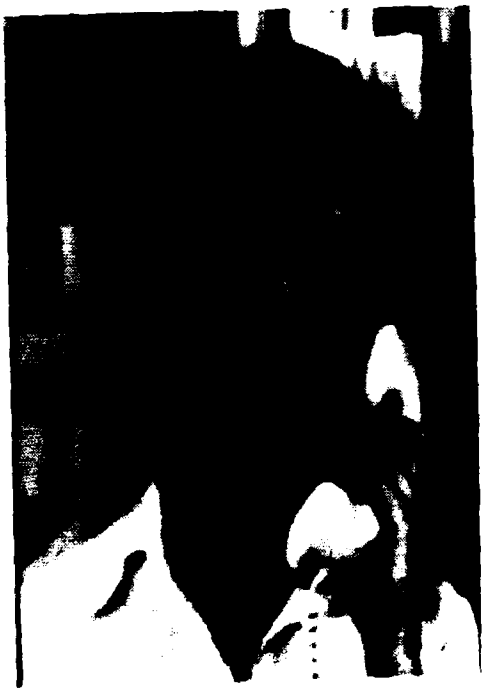


(f) Output Image 2

Figure 20 Continued.

(g) Output Image 3 (.392)

(h) Output Image 4 (.564)



(i) Output Image 5 (.893)

(j) Output Image 6 (1.22)

Figures 21 a through f are the originals for the 'SLOW PHONE' image sequence while 22 a through e show the outputs and the associated bit rates.

(a) Input Image 1.

(b) Input Image 2

(c) Input Image 3.

(d) Input Image 4.

Figure 21 Continued.

(e) Input Image 5.

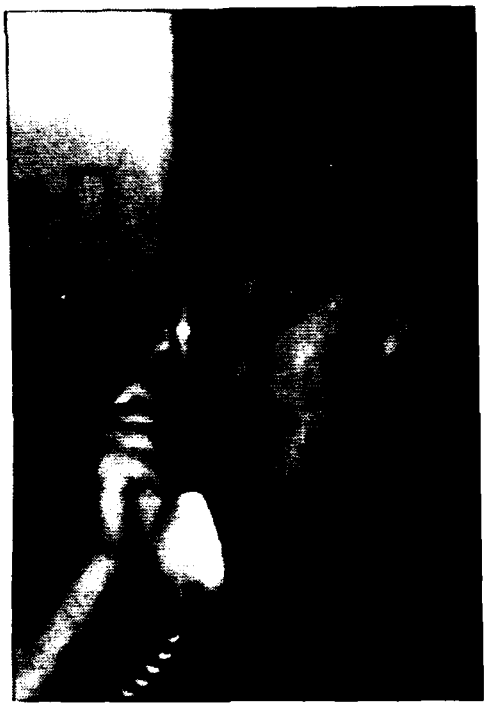
(f) Input Image 6.

Figure 22. Output Image Sequence

(a) Output Image 2 (.453)

(b) Output Image 3 (.693)

Figure 22 Continued.



(c) Output Image 4 (.525)

(d) Output Image 5 (.481)

(e) Output Image 6 (.451)

Finally figures 23 a through f give the original for the 'PLANE' sequence and figures 24 a through e give the outputs with the associated bit rates.

(a) Input Image 1

(b) Input Image 2

(c) Input Image 3

(d) Input Image 4

Figure 23 Continued

(e) Input Image 5

(f) Input Image 6

Figure 24 Output Image Sequence

(a) Output Image 2 (.595)

(b) Output Image 3 (.489)

Figure 24 Continued

(c) Output Image 4 (.689)

(d) Output Image 5 (.598)

(e) Output Image 6 (.724)

REFERENCES

1. Agin, G.J., "Computer Vision Systems for Industrial Inspection and Assembly," IEEE Computer, Volume 13, Number 5, May 1980, pp. 11-20.

2. Anderson, R.F., "TV-Compatible Forward Looking Infrared," Optical Engineering, Volume 13, Number 4, July-August 1974, pp. 335-338.

3. Anderson, B.D.O., and Moore, J.B., Optimal Filtering, Prentice-Hall Inc., 1979.

4. Andrews, H.C., and Hunt, B.R., Digital Image Restoration, Prentice-Hall Inc., 1977.

5. Ayala, I.L., Orton, D.A., Larson, J.B., and Elliot, D.F., "Moving Target Tracking Using Symbolic Registration," IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume PAMI-4, Number 5, September 1982, pp. 515-520.

6. Barakat, Richard, "Application of the Sampling Theorem to Optical Diffraction Theory," Journal of the Optical Society of America, Volume 54, Number 7, July 1964, pp. 920-930.

7. Barnes, Casper W., "Object Restoration in a Diffraction-Limited Imaging System," Journal of the Optical Society of America, Volume 56, Number 5, May 1966, pp. 575-576.

8. Beall, W.H., "Statistical Analysis of Degradation in Scanned Image Systems," Journal of the Optical Society of America, Volume 54, Number 8, August 1964, pp. 992-997.

9. Bode, H.W., and Shannon, C.E., "A Simplified Derivation of Linear Least Squares Smoothing and Prediction Theory," Proceedings of the IRE, Volume 38, pp. 417-425, April 1950.

10. Bowling, Carl D., and Jones, R.A., "Displacement Estimation by Prediction Coefficient Energy Concentration," IEEE Globall Telecommunications Conference, GLOBECOM'82, Volume 1, pp. B6.4.1-B6.4.4.

11. Budrikis, Z.L., "Visual Fidelity Criterion and Modeling," Proceedings of the IEEE, Volume 60, Number 7, July 1972, pp. 771-779.

12. Castleman, K.R., Digital Image Processing, Prentice-Hall Inc., 1979.

13. Chen, Wen-Hsiung, and Smith, C. Harrision, "Adaptive Coding of Monochrome and Color Images," IEEE Transactions on Communications, Volume COM-25, Number 11, November 1977, pp. 1285-1292.

14. Cooper, G.R., and McGillem, C.D., Probabilistic Methods of Signal and System Analysis, Holt, Rinehart, and Winston Inc., 1971.

15. Danker, A.J., and Rosenfeld, A., "Blob Detection by Relaxation," IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume PAMI-3, Number 1, January 1981, pp. 79-92.

16. Daut, D.G., Fries, R.W., and Modestino, J.W., "Two-Dimensional DPCM Image Coding Based on an Assumed Stochastic Image Model," IEEE Transactions on Communications, Volume COM-29, Number 9, September 1981, pp. 1365-1374.

17. Dukhovich, Isaac J., and O'Neal, J.B. Jr., "A Three-Dimensional Spatial Non-Linear Predictor for Television," IEEE Transactions on Communications, Volume COM-26, Number 5, May 1978, pp. 578-583.

18. Ferrie, Frank P., Levine, Martin D., and Zucker, Steven W., "Cell Tracking: A Modeling and Minimization Approach," IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume PAMI-4, Number 3, May 1982, pp. 277-291.

19. Forgus, R.H., and Melamed, L.E., Perception: A Cognitive Stage Approach, McGraw-Hill Book Company, 1976.

20. Gallager, R.G., Information Theory and Reliable Communication, John Wiley and Sons Inc., 1968.

21. Gilbert, Alton L., Giles, Michael K., Flachs, Gerald M., Rogers, Robert B., and U, Yee Hsun, "A Real-Time Video Tracking System," IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume PAMI-2, Number 1, January 1980, pp. 47-56.

22. Gonzalez, R.C., and Wintz, P., Digital Image Processing, Addison-Wesley Publishing Company Inc., 1979.

23. Goyal, Shri K., and O'Neal, J.B. Jr., "Entropy Coded Differential Pulse-Code Modulation Systems for Television," IEEE Transactions on Communications, Volume COM-23, Number 6, June 1975, pp. 660-666.

24. Granrath, Douglas J., "The Role of Human Vusual Models in Image Processing," Proceedings of the IEEE, Volume 69, Number 5, May 1981, pp. 552-561.

25. Habibi, Ali, and Wintz, Paul A., "Image Coding by Linear Transformation and Block Quantization," IEEE Transactions on Communications, Volume COM-19, Number 1, February 1971, pp. 50-62.

26. Habibi, Ali, "Hybrid Coding of Pictorial Data," IEEE Transactions on Communications, Volume COM-22, Number 5, May 1974, pp. 614-624.

27. Habibi, Ali, "Survey of Adaptive Image Coding Techniques," IEEE Transactions on Communications, Volume COM-25, Number 11, November 1977, pp. 1275-1284.

28. Habibi, Ali, "An Adaptive Strategy for Hybrid Image Coding," IEEE Transactions on Communications, Volume COM-29, Number 12, December 1981, pp. 1736-1740.

29. Harris, J.L., "Diffraction and Resolving Power," Journal of the Optical Society of America, Volume 54, Number 7, July 1964, pp. 931-936.

30. Harris, J.L. Sr., "Image Evaluation and Restoration," Journal of the Optical Society of America, Volume 56, Number 5, May 1966, pp. 569-574.

31. Haskell, Barry G., Gordon, Pat L., Schmidt, Robert L., and Scattaglia, James V., "Interframe Coding of 525-Line Monochrome Television at 1.5 Mbits/s," IEEE Transactions on Communications, Volume COM-25, Number 11, November 1977, pp. 1339-1348.

32. Healy, Donald J., Mitchell, O.R., "Digital Video Bandwidth Compression Using Block Truncation Coding," IEEE Transactions on Communications, Volume COM-29, Number 12, December 1981, pp. 1809-1817.

33. Huang, T.S., Picture Processing and Digital Filtering, Springer-Verlag, 1975.

34. Jain, R., "Extraction of Motion Information from Peripheral Processes," IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume PAMI-3, Number 5, September 1981, pp. 489-503.

35. Jain, Anil K., and Angel, Edward, "Image Restoration, Modelling, and Reduction of Dimensionality," IEEE Transactions on Computers, Volume C-23, Number 5, May 1974, pp. 470-476.

36. Jain, Jaswant R., "Displacement Measurement and Its Application in Interframe Image Coding," IEEE Transactions on Communications, Volume COM-29, Number 12, December 1981 pp. 1799-1808.

37. Jain, A.K., "Advances in Mathematical Models for Image Processing," Proceedings of the IEEE, Volume 69, Number 5, May 1981, pp. 502-528.

38. Jain, Ramesh, and Haynes, Susan, "Imprecision in Computer Vision," IEEE Computer, Volume 15, Number 8, August 1982, pp. 39-48.

39. Jarvis, J.E., "Visual Inspection Automation," IEEE Computer, Volume 13, Number 5, May 1980, pp. 32-38.

40. Jarvis, R.A., "A Computer Vision and Robotics Laboratory," IEEE Computer, Volume 15, Number 6, June 1982, pp. 8-22.

41. Jones, R.A., "Adaptive Hybrid Picture Coding," SPIE, Vol.87, Advances in Images Processing Techniques, 1976, pp. 247-255.

42. Jones, R.A., "Adaptive Hybrid Picture Coding," Technical Proposal D180-19071-1, Boeing Aerospace Company, Research & Engineering Division, September 1976.

43. Jones, R.A., "Adaptive Hybrid Picture Coding," Interim Scientific Report, Department of Electrical Engineering, University of Arkansas, January 1979.

44. Jones, R.A., "Adaptive Hybrid Picture Coding," Second Interim Scientific Report, Department of Electrical Engineering, University of Arkansas, November 1979.

45. Kailath, T., "An Innovations Approach to Least-Squares Estimation. Part I: Linear Filtering in Additive White Noise," IEEE Transactions on Automatic Control, Volume AC-13, Number 6, December 1968, pp. 646-655.

46. Kailath, T. and Frost, P., "An Innovations Approach to Least-Squares Estimation. Part II: Linear Smoothing in Additive White Noise," IEEE Transactions on Automatic Control, Volume AC-13, Number 6, December 1968, pp. 655-660.

47. Kogo, Toshio, Iijima, Y., and Iinuma, K., "Statistical Performance Analysis of an Interframe Encoder for Broadcast Television signals," IEEE Transactions on Communications, Volume COM-29, Number 12, December 1981, pp. 1868-1876.

48. Kuo, B.C., Digital Control Systems, Holt, Rinehart and Winston Inc., New York, 1980.

49. Legters, George R. Jr., and Young, Tzay Y., "A Mathematical Model for Computer Image Tracking," IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume PAMI-4, Number 6, November 1982, pp. 583-594.

50. Levy, Armand Jacques, "Wiener Estimation for Inversion of Linear Operators and Superresolution," Proceedings IEEE Computer Society Conference on Pattern Recognition and Image Processing, 1981, pp. 433-436.

51. Liew, Chong Kiew, "Inequality Constrained Least-Squares Estimation," Journal of the American Statistical Association, Volume 71, Number 355, September 1976, Theory and Methods Section, pp. 746-751.

52. Liew, Chong K., and Shim, Jae K., "A Computer Program for Inequality Constrained Least-Squares Estimation," Econometrica, Volume 46, Number 1, January 1978, pp. 237.

53. Marchand, E.W., "Derivation of the Point Spread Function from the Line Spread Function," Journal of the Optical Society of America, Volume 54, Number 7, July 1964, pp. 915-919.

54. Marks, Robert J. II, "Gerchberg's Extrapolation Algorithm in Two Dimensions," Applied Optics, Volume 20, Number 10, May 15, 1981, pp. 1815-1820.

55. Marks, Robert J. II, and Smith, Michael J., "Closed-form Object Restoration from Limited Spatial and Spectral Information," Optics Letters, Volume 6, Number 11, November 1981, pp. 522-524.

56. Maxemchuk, N.F., and Stuller J.A., "An Intraframe Codec Based Upon Non Stationary Image Model," Bell System Technical Journal, Volume 58, Number 6 Part II, July-August 1979, pp. 1395-1412.

57. Maxemchuk, N.F., and Stuller, J.A., "Reduction of Transmission Error Propagation in Adaptively Predicted, DPCM Encoded Pictures," Bell System Technical Journal, Volume 58, Number 6 Part II, July-Sugust 1979, pp. 1413-1423.

58. Meiri, A. Zvi, and Yedilevich, E., "A Pinned Sine Transform Image Coder," IEEE Transactions on Communications, Volume COM-29, Number 12, December 1981, pp. 1728-1735.

59. Melsa, James L., and Sage, Andrew P., Estimation Theory with Applications to Communications and Control, McGraw-Hill Book Company, 1971.

60. Melsa, J.L., and Cohn, D.L., Decision and Estimation Theory, McGraw-Hill Book Company, 1978.

61. Meyers, W., "Industry Begins to Use Visual Pattern Recognition," IEEE Computer, Vol.13, Number 5, May 1980, pp. 21-31.

62. Mitchell, Owen R., and Tavatavai, A.J., "Channel Error Recovery for Transform Image Coding," IEEE Transactions on Communications, Volume COM-29, Number 12, December 1981, pp. 1754-1762.

63. Mix, Dwight, F., Random Signal Analysis, Addison-Wesley Publishing Company, 1969.

64. Modestino, James W., and Bhaskaran, V., "Robust Two-Dimensional Tree Encoding of Images," IEEE Transactions on Communications, Volume COM-29, Number 12, December 1981, pp. 1786-1798.

65. Modestino, James W., Daut, D.G., and Vickers, A.L., "Combined Source-Channel Coding of Images Using Block Cosine Transform," IEEE Transactions on Communications, Volume COM-29, Number 9, September 1981, pp. 1261-1274.

66. Modestino, James W., Bhaskaran, V., and Anderson, J.B., "Tree Encoding of Images in the Presence of Channel Errors,", IEEE Transactions on Information Theory, Volume IT-27, Number 6, November 1981, pp. 677-697.

67. Mohanty, N.C., "Computer Tracking of Moving Point Targets in Space," IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume Pami-3, Number 5, September 1981, pp. 606-611.

68. Mohwinkel, C., and Kurz, L., "Computer Picture Processing and Enhancement by Localized Operations," Computer Graphics and Image Processing, Volume 5, 1976, pp. 401-424.

69. Natarajan, T. Raj, Ahmed, Nasir, "On Interframe Transform Coding," IEEE Transactions on Communications, Volume COM-25, Number 11, November 1977, pp. 1323-1329.

70. Netravali, A.N., and Robbins, J.D., "Motion-Compensated Television Coding: Part I," The Bell System Technical Journal, Volume 58, Number 3, March 1979, pp. 631-669.

71. Netravali, A.N., and Stuller, J.A., "Motion-Compensated Transform Coding," Bell System Technical Journal, Volume 58, Number 7, September 1979, pp. 1703-1718.

72. Netravali, Arun N., and Limb, John O., "Picture Coding: A Review," Proceedings of the IEEE, Volume 68, Number 3, March 1980, pp. 366-406.

73. Netravali, A.N., and Bowen, E.G., Bell System Technical Journal, Volume 61, Number 6, July-August 1982, pp. 969-979.

74. Ngan, K.N., "Adaptive Transform Coding of Video Signals," IEE Proceedings, Volume 129, Part F, Number 1, February 1982, pp. 28-40.

75. O'Brien, Kathryn B., "Automatic Optical Design of Desired Image Distributions Using Orthogonal Constraints," Journal of the Optical Society of America, Volume 54, Number 10, October 1964, pp. 1252-1255.

76. Oppenheim, Alan V., Schafer, Ronald W., and Stockham, Thomas G., "Nonlinear Filtering of Multiplied and Convolved signals," Proceedings of the IEEE, Volume 56, Number 8, August 1968, pp. 1264-1291.

77. O'Rourke, J., "Motion Detection Using Hough Techniques," IEEE Computer Society Conference on Pattern Recognition and Image Processing, 1981, pp. 82-87.

78. Ostrem, J.S., and Falconer, D.G., "A Differential Operator Technique for Restoring Degraded Signals and Images," IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume PAMI-3, Number 3, May 1981, pp. 278-284.

79. Pack, C.D., and Whitaker, B.A., "Kalman Filter Models for Network Forecasting," Bell System Technical Journal, Volume 61, Number 1, January 1982, pp. 1-15.

80. Pal, S.K., and King, R.A., "Image Enhancement Using Smoothing with Fuzzy Sets," IEEE Transactions on Systems, Man, and Cybernetics, Volume SMC-11, Number 7, July 1981, pp. 494-501.

81. Papoulis, A., Probability, Random Variables, and Stochastic Processes, McGraw-Hill Book Company, 1965.

82. Peacock, K.L., and Treitel, Sven, "Predictive Deconvolution: Theory and Practice," Geophysics, Volume 34, Number 2, April 1969, pp. 155-169.

83. Peleg, Shmuel, "A New Probabilistic Relaxation Scheme," IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume PAMI-2, Number 4, July 1980, pp. 362-369.

84. Pirsh, P., "Adaptive Intra-Interframe DPCM Coder," Bell System Technical Journal, Volume 61, Number 5, May-June 1982, pp. 747-764.

85. Porter, G.B. and Mundy, J.L., "Visual Inspection System Design," IEEE Computer, Volume 13, Number 5, May 1980, pp. 40-48.

86. Pratt, W.K., Digital Image Processing, Wiley Interscience, 1978.

87. Price, C., Snyder, W., and Rajala, S., "Computer Tracking of Moving Objects Using a Fourier-Domain Filter Based on a Model of the Human Visual System," IEEE Computer Society Conference on Pattern Recognition and Image Processing, 1981, pp. 98-102.

88. Rashid, H.U., and Jones, R.A., "An Adaptive Estimation Approach to Estimation Displacement in Time Varying Images," SPIE 23rd Internation Symposium, August 1979, San Diego California.

89. Rashid, H.U., "An Innovations Approach to Displacement Estimation and Image Analysis in Time Varying Images," PHD Dissertation, Department of Electrical Engineering, University of Arkansas, August 1980.

90. Robbins, John D., and Netravali, Arun N., "Interframe Telivision Coding Using Movement Compensation," Internation Conference on Communications Record, ICC'79, Volume 2, pp.23.4.1-23.4.5.

91. Robbins, J.D., and Netravali, A.N., "Spatial Subsampling in Motion-Compensated Television Coders," Bell System Technical Journal, Volume 61, Number 8, October 1982, pp. 1895-1910.

92. Roese, John A., Pratt, William K., and Robinson, Guner S., "Interframe Cosine Transform Image Coding," IEEE Transactions on Communications, Volume COM-25, Number 11, November 1977, pp. 1329-1339.

93. Rosenfeld, A., Picture Processing by Computer, Academic Press Inc., 1969.

94. Rosenfeld, Azriel, and Kak, Avinash, Digital Picture Processing, Academic Press Inc., 1976.

95. Sabri, M. Shaker, and Steenaart, Willem, "An Approach to Band-limited Signal Extrapolation: The Extrapolation Matrix," IEEE Transactions on Circuits and Systems, Volume 25, Number 2, February 1978, pp. 74-78.

96. Sabri, Shaker "Movement-Compensated Interframe Prediction for NTSC Colour TV Signals," Paper Presented at the Nato Advanced Study Institute on Image Sequence Processing and Dynamic Scene Analysis, Braunlage, West Germany, June 1982.

97. Sakai, Hajime, and Vanasse, George A, "Direct Determination of the Transfer Function of an Infrared Spectrometer," Journal of the Optical Society of America, Volume 56, Number 3, March 1966, pp. 357-362.

98. Schalkoff, R.J., and McVey, E.S., "A Model and Tracking Algorithm for a Class of Video Targets," IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume PAMI-4, Number 1, January 1982, pp. 2-10.

99. Shapiro, Linda G., and Haralick, Robert M., "Organization of Relational Models for Scene Analysis," IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume PAMI-4, Number 6, November 1982, pp. 595-602.

100. Smith, David K., and Marks, Robert J. II, "Closed Form Bandlimited Image Extrapolation," Applied Optics, Volume 20, Number 14, July 15 1981, pp. 2476-2483.

101. Stark H., Cahana, D., and Webb, H., "Restoration of Arbitrary Finite-Energy Optical Objects from Limited Spatial and Spectral Information," Optical Society of America, Volume 71, Number 6, June 1981, pp. 635-642.

102. Stockham, Thomas G., "Image Processing in the Context of a Visual Model," Proceedings of the IEEE, Volume 60, Number 7, July 1972, pp. 828-842.

103. Stuller, J.A., Netravali, A.N., and Robbins, J.D., "Interframe Television Coding Using Gain and Displacement Compensation," Bell System Technical Journal, Volume 59, Number 7, September 1980, pp. 1227-1240.

104. Tou, J.T., and Gonzalez, R.C., Pattern Recognition Principles, Addison-Wesley Publishing Company, Reading, Mass., 1974.

105. Webb, J.A., and Aggarwal, J.K., "Visual Interpretation of the Motion of Objects in Space," IEEE Computer Society Conference on Pattern Recognition and Image Processing, 1981, pp. 516-521.

106. Williams, R.A., and Chang, W.S.C., "Resolution and Noise in Fourier-Transform Spectroscopy," Journal of the Optical Society of America, Volume 56, Number 2, February 1966, pp. 167-170.

107. Winston, Patrick Henry, The Psychology of Computer Vision, McGraw-Hill Book Company, 1975.

108. Wintz, Paul A., "Transform Picture Coding," Proceedings of the IEEE, Volume 60, Number 7, July 1972, pp. 809-820.

109. Youla, Dante C., "Generalized Image Restoration by the Method of Alternating Orthogonal Projections," IEEE Transactions on Circuits and Systems, Volume 25, Number 9, September 1978, pp. 694-702.

110. Zadeh, L.A., and Ragazzini, "An Extension of Wiener's Theory of Prediction," Journal of Applied Physics, Volume 21, pp. 545-655, July 1950.

111. Zschunke, Willmut, "DPCM Picture Coding with Adaptive Prediction," IEEE Transactions on Communications, Volume COM-25, Number 11, November 1977, pp. 1295-1302.

PART TWO

MACHINE RECOGNITION OF PARTIAL SHAPES

USING

FEATURE VECTORS DEFINED ON SHAPE SPACE

MACHINE RECOGNITION OF PARTIAL SHAPES USING FEATURE VECTORS DEFINED ON SHAPE SPACE

INTRODUCTION

In this paper the problem of determining whether a partial shape belongs to a more complex whole shape is addressed. A partial shape occurs when a shape is partially obscured by another shape, or object, or is partly in the field of view.

At present the methods of recognizing shapes [5], [15], [21], [22], can be categorized as either global or local in nature. Within the class of global shape analysis algorithms there are two categories that under certain circumstances possess the ablity to recognize complete or whole shapes independent of size, rotation, or location. These are the Fourier descriptors methods and the Syntactic, or Graphical, methods. The Fourier descriptors based algorithm perform satisfactorily on complete shapes. The Fourier coefficients extracted are indeed independent of size, rotation, and location when the shapes are complete. However this method does not perform satisfactorily and in fact fails entirely when the class of shapes is allowed to include incomplete or partial shapes. An example is presented in Section II that demonstrates that Fourier descriptors method fails to work

on incomplete shapes. The results of the experiment are discussed in order to point out specifically why the algorithm cannot perform satisfactorily on partial shapes.

The other class of global methods, namely the Syntactic methods [4] have restricted use in recognition of shapes because these algorithms tacitly assume a priori that the shapes have been identified by their parts. These algorithms then investigate the relationship between the various parts of the shape. A human shape, for instance has a hand or a face at some definite orientaion and location with respect to each other.

The Local category of shape analysis algorithm [21] uses 'curvature' as a criterion for detecting the peaks and valleys of a shape. These peaks and valleys are called the local shape descriptors. This shape comparison algorithm is not independent of rotation. In Section II specific examples are given that demonstrate that the present comparison type local shape analysis algorithms are not independent of rotation. The concept of curvature is then presented from the point of view of differential geometry for determining why the local shape descriptors are not independent of rotation.

In Section III, several general concepts from the allometric disciplines [11], are combined with some entirely new concepts concerning shapes for the purpose of providing the foundation for a new approach to defining shape as

on incomplete shapes. The results of the experiment are discussed in order to point out specifically why the algorithm cannot perform satisfactorily on partial shapes.

The other class of global methods, namely the Syntactic methods [4] have restricted use in recognition of shapes because these algorithms tacitly assume a priori that the shapes have been identified by their parts. These algorithms then investigate the relationship between the various parts of the shape. A human shape, for instance has a hand or a face at some definite orientaion and location with respect to each other.

The Local category of shape analysis algorithm [21] uses 'curvature' as a criterion for detecting the peaks and valleys of a shape. These peaks and valleys are called the local shape descriptors. This shape comparison algorithm is not independent of rotation. In Section II specific examples are given that demonstrate that the present comparison type local shape analysis algorithms are not independent of rotation. The concept of curvature is then presented from the point of view of differential geometry for determining why the local shape descriptors are not independent of rotation.

In Section III, several general concepts from the allometric disciplines [11], are combined with some entirely new concepts concerning shapes for the purpose of providing the foundation for a new approach to defining shape as

vectors in the appropriate space. The properties of this space are stated definitively, after the concept of size variable has been solidified. This vector space is called shape space. Two theorems useful in the partial shape recognition problem are stated and proved utilizing shape space properties.

In Section IV, some basic definitions regarding critical points are presented. These definitions are used in Section V where one of the principal results of this paper is presented. Specifically, a new procedure of determining the critical points of a shape is described. This procedure is named the Adaptive Line of Sight method. In the Adaptive Line of Sight method, the critical point determination in based on a set of coordinate axes that are dependent on the shape being examined. Examples are given that demonstrate that the procedure produces critical points that are independent of rotaion, size, displacement, and correspond closely to those produced by normal human cognitive process.

It is demonstrated in Section VI that measurements between a set of critical points that are determined by using the Adaptive Line of Sight method can be used to define feature vectors for shapes. These feature vectors remain the same whether the shape is a partial shape or a more complex whole shape. Moreover, these feature vectors are independent of size, rotation and displacement since they are derived from a set of critcal points that are

independent of the same quantities. A procedure for comparing the feature vectors of a set of shapes is described The comparison procedure is based on a Syntactic method which will point out whether one shape is part of a more complex whole shape, or whether the shapes are totally dissimilar.

SECTION II

PRELIMINARY THEORY AND DEFINITIONS

There exist two types of Fourier descriptors. The first type of descriptors, used by Zahn and Roskies [22], have been called descriptors S_n by Pavlidis [13]. In the method of descriptors S_n the shape is represented by the continuous function,

$$a(t(k),k) = \phi(k) + t(k) \qquad (2.1)$$

where $t(k) = 2\pi\rho(k)/L$

$\rho(k)$ = arc length between the starting point and the k th point on the curve.

$\phi(k)$ = net amount of angular change between the starting point and the k th point on the curve.

L = the perimeter of the curve

The descriptors S_n for a continuous shape are then defined as,

$$S_n = \frac{1}{2\pi}\int_0^{2\pi} a(t(k), k)\exp(-j2\pi nt)\,dt \qquad (2.2)$$

These descriptors exhibit some notable shortcomings. Among these are 1) the property of the closure of the curve is not preserved 2) simple shapes such as squares and triangles cannot be distinguished from one another when only the ver-

tices are given. The second type of Fourier descriptors, namely the descriptors T_n [13], [15], [17] exhibit characteristics that are superior to the descriptors S_n in the sense that the reconstruction of the shape from a finite set of coefficients leads to a closed curve. Also the convergence properties are superior.

For the descriptors T_n the shape data is represented in the complex form,

$$u(\rho) = x(\rho) + j\, y(\rho) \tag{2.3}$$

where $(x(\rho)\ ,\ y(\rho))$ are the coordinates of the point on the curve and ρ is the arc length *from the defined starting* point. The descriptors T_n are then given by,

$$T_n = \frac{1}{L}\int_0^L u(\rho)\exp(-j2\pi n\rho/L)\,dl \tag{2.4}$$

The Fourier descriptors based algorithm normalizes for position by setting T_o to zero . Normalization for scale, rotation, and starting point of a contour is achieved by multiplying the n th coefficient by s exp(j (ϕ +nα)). The parameter s scales the shape to the normalized size, the parameter (ϕ + nα) rotates the contour to the normalized position. The normalization is such that the coefficients T_{+1} and T_{-1} are pure imaginary numbers and their sum has magnitude one.

It is easy to see why such a normalization is necessary for discrete data. In this case, the Fourier descriptors are given by,

$$T(n) = \frac{1}{N} \sum_{k=0}^{N-1} (x(k) + j\,y(k)) \exp(-j2\pi nk/N) \qquad (2.5)$$

Multiplying (2.5) by s exp$-j(\phi + n\alpha)$ the normalized coefficients, $T_N(n)$ are then given as,

$$T_N(n) = T(n)\, s\, \exp j(\phi + n\alpha) \qquad (2.6)$$

Next imposing the requirement on (2.6), that the coefficient be purely imaginary at $n = \pm 1$ leads to the condition that,

$$\frac{2\pi k}{N} - (\phi + \alpha) = 2m + \left(-\frac{2\pi k}{N} - (\phi - \alpha)\right) \qquad (2.7)$$

where m is an integer.

Setting the real part to zero by appropriate choices of ϕ and α is equavalent to normalizing for rotation and starting point of the contour. As a matter of fact, the two equations could have been set equal to any constant.

The imaginary parts of the normalized coefficient at $n = \pm 1$ are,

$$IM(T_N(1)) = \frac{s}{N} \sum_{k=0}^{N-1} \left(-x(k)\sin\left(\frac{2\pi k}{N} - (\phi + \alpha)\right) + y(k)\cos\left(\frac{2\pi k}{N} - (\phi + \alpha)\right)\right)$$

$$IM(T_N(-1)) = \frac{s}{N} \sum_{k=0}^{N-1} \left(-x(k)\sin\left(\frac{2\pi k}{N} - (\phi - \alpha)\right) + y(k)\cos\left(\frac{2\pi k}{N} - (\phi - \alpha)\right)\right)$$

Summing and equating the magnitude, of the normalized coefficients at ± 1, to one, yields,

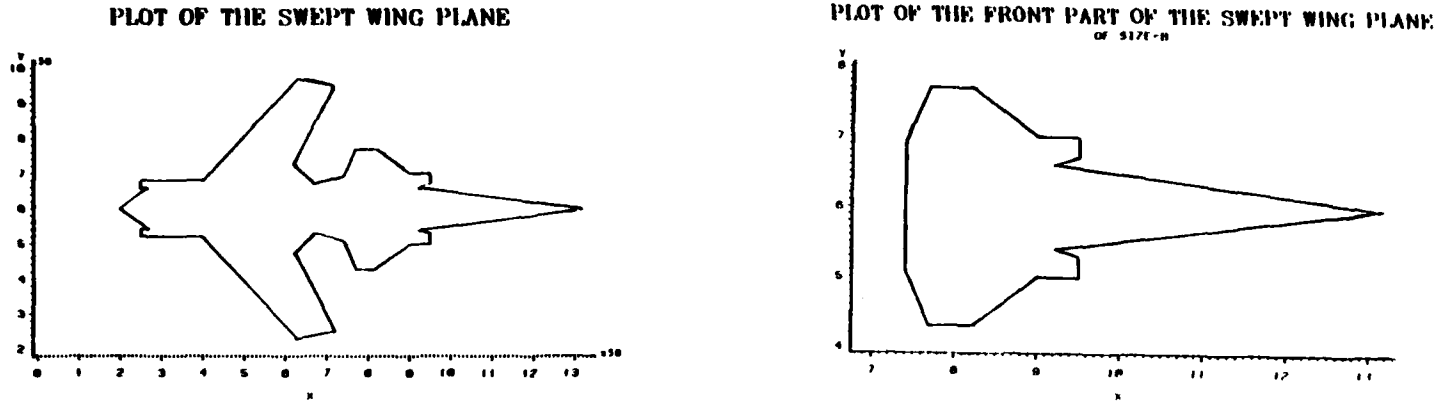


FIG. 1. SWEPT-WING PLANE SHAPES.

$$s = \{(\sum_{k=0}^{N-1} (x(k)\sin(A+B) - \sin(A-C)) + (y(k)(\cos(A-C)+\cos(A-B)))/N\}^{-1} \quad (2.8)$$

where $A = 2\pi k/N$

$B = \phi + \alpha$

$C = \phi - \alpha$

Another function [16] which has also been used for normalization is the standard deviation of the data,

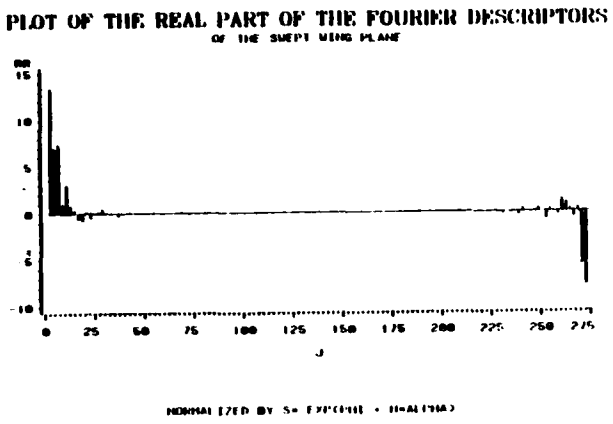
$$\sigma = \left[\frac{N}{\sum_{k=0}^{N-1} (x(k)-\bar{x}) + (y(k)-\bar{y})}\right]^{\frac{1}{2}} \quad (2.9)$$

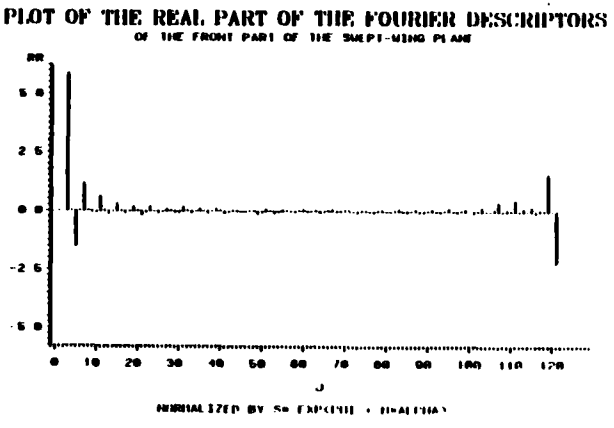
where $\bar{x} = \frac{1}{N} \sum_{k=0}^{N-1} x(k)$

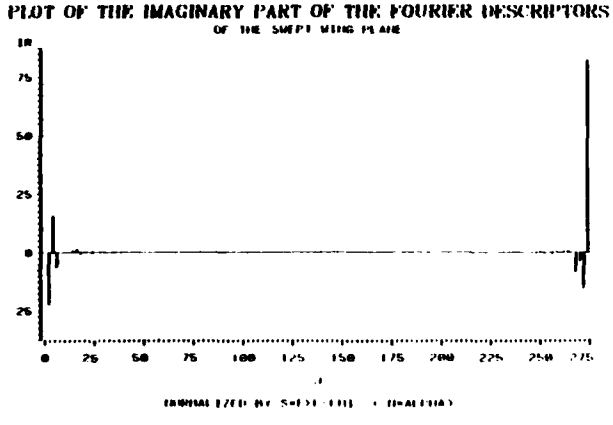
$\bar{y} = \frac{1}{N} \sum_{k=0}^{N-1} y(k)$

It is apparent that both s and σ are linear homogeneous functions of the data points.

In a real shape recognition application it is not known a priori whether the shape under examination is a part of a more complex shape or a shape in its own right. The above descriptors in their present form are not capable of recognizing that a partial shape may be part of a more complex shape. We present an example to demonstrate the validity of this contention. The shapes used are shown in Fig. 1. Figure 1-a is a complete shape, namely a swept wing plane, while Figure 1-b is a partial shape, namely the front part of the







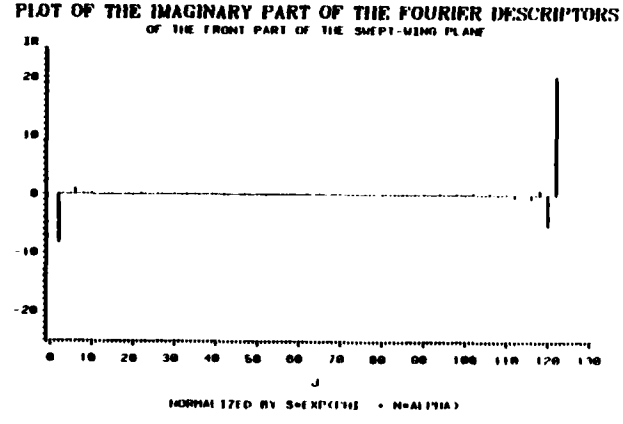


FIG. 2. FOURIER DESCRIPTORS CALCULATED USING THE PARAMETER s.

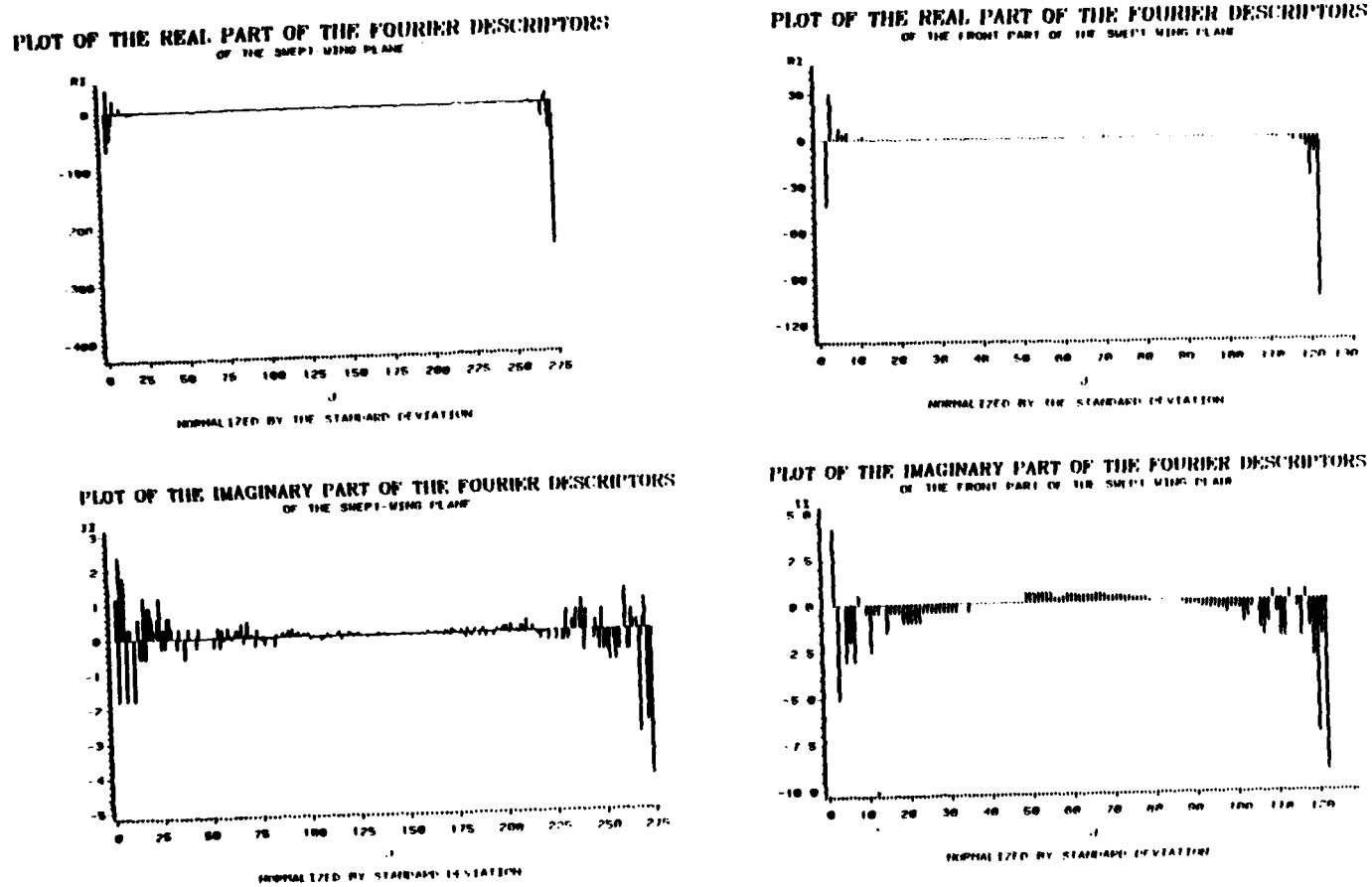


FIG. 3. FOURIER DESCRIPTORS CALCULATED USING THE PARAMETER σ.

plane. The plots of the real and imaginary part of the normalized Fourier descriptors obtained by using (2.8), for the complete shape are shown in Fig. 2-a. The corresponding plots for the partial shape are are shown in Fig. 2-b. Similar data obtained using (2.9) is shown in Fig. 3. It is apparent that comparing the two sets of data yields nothing more that a statement that the two shapes are dissimilar. Two explanations for this are 1) the parameters s and σ are not independent of the shape and 2) the Fourier descriptors method compares the 'frequencies' of the shapes. The frequencies of the partial shape are not the same as the frequencies of the complete shape.

The main result of this paper is the description of an approach for determining if a partial shape belongs to a more complex whole shape. This is detailed in section III IV and V.

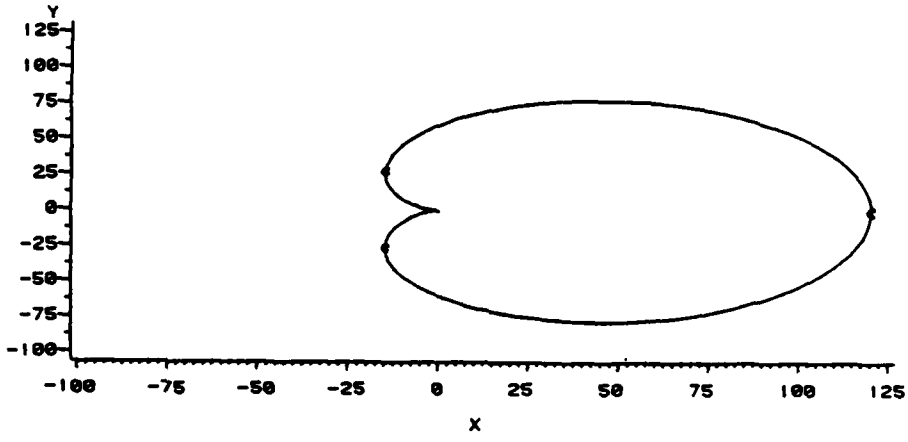
II-B LOCAL SHAPE DESCRIPTORS

The Local Shape descriptors method has been used by Wallace, Mitchell and Fukunaga [21], for comparing shapes stored in a library. In this method, the angle function is defined as,

$$s(k)=\arctan((y(k)-y(k-1))/(x(k)-x(k-1)) \qquad (2.10)$$

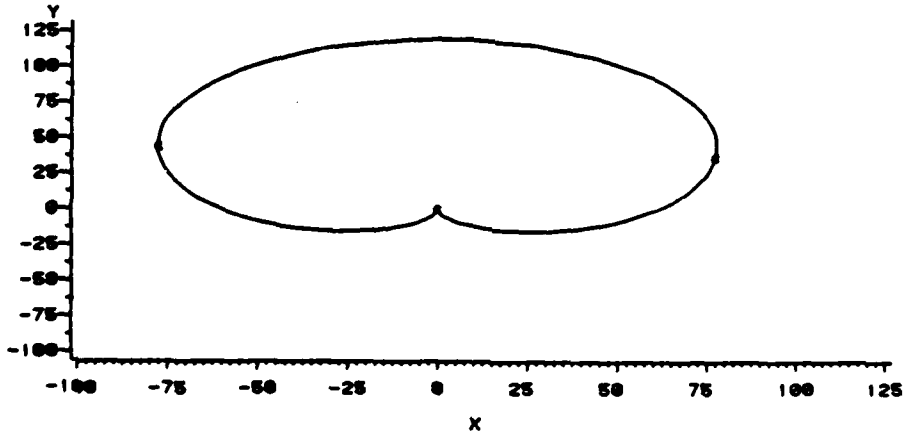
Curvature is then defined as the derivative of the angle function, that is s'(k). The peaks and valleys in s'(k)

CARDIOID OF SIZE=A



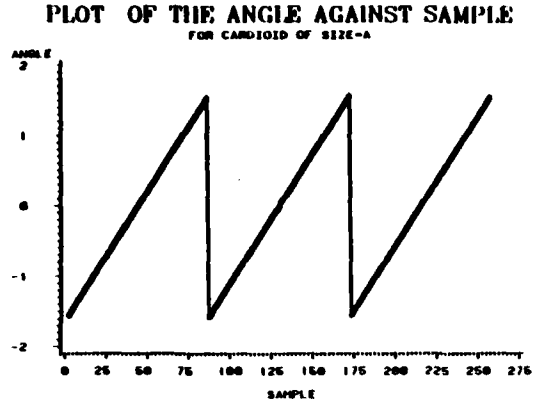
C=CRITICAL POINTS FOUND BY TAKING THE DERIVATIVE OF THE ANGLE

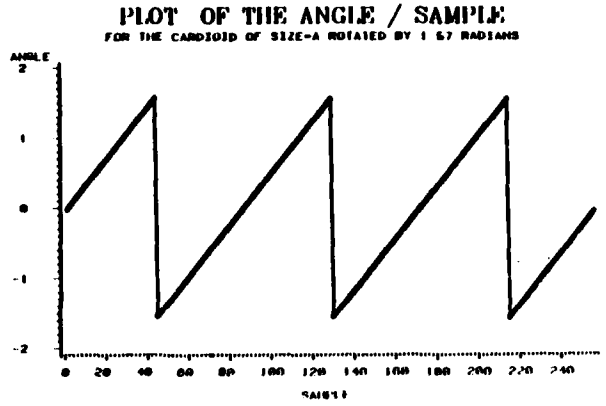
CARDIOID OF SIZE=A ROTATED BY 1.57 RADIANS

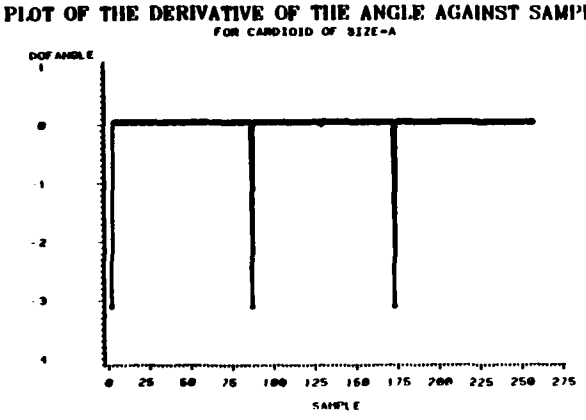


C=CRITICAL POINTS FOUND BY TAKING THE DERIVATIVE OF THE ANGLE

FIG. 4. CRITICAL POINTS OF THE CARDIOID SHAPE.







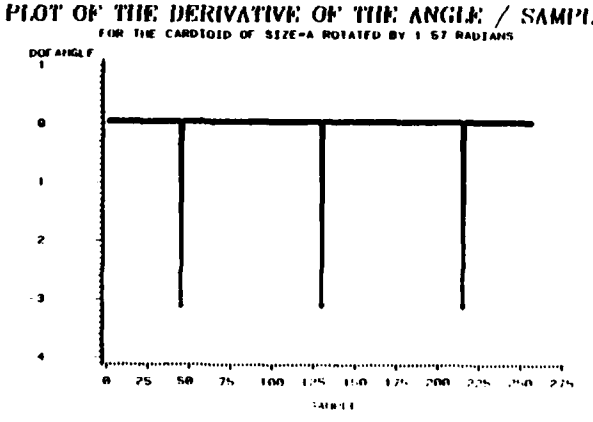


FIG. 5. ANGLES AND DERIVATIVES OF ANGLES FOR THE CARDIOID SHAPE.

are used for finding the peaks and valleys of the curve. Each local shape descriptor then consists of two adjacent peaks and a valley (alternate angles and a distance).

Observe that (2.10) is a nonlinear transformation on the data, also, the function is not finite at $\pm$ 90°. In general, the function may be used to locate peaks where 1) the angle between any three adjacent points is acute 2) the shape becomes parallel to the y-axis. This is equivalent to places where the angle becomes $\pm \frac{\pi}{2}$ radians with respect to the x-axis. To illustrate this point the cardioid shown in Fig. 4-a was generated sampling at constant intervals of $(2\pi/256)$ radians and not the arc length, so that an acute angle did not occur between any three adjacent points. The cardioid was then rotated by 1.57 radians. The corresponding plots for the angle, and the derivative of the angle function, and the peaks obtained using this method are shown in Fig. 5 and Fig. 6, for both cases. Note peaks obtained by this method are not unique and are dependent on the rotation.

Curvature is known [7] to be a property of a curve independent of its orientation in space. Despite the fact that this orientation independent quantity was used to determine peaks, the peaks were not independent of rotation. Clarification of this apparent contradiction requires the concepts of curvature and torsion of curves in space.

CURVES IN THREE DIMENSIONAL SPACE

The problem of shape recognition is analogous to recognition of curves in space. Therefore, well known concepts and theorems from differential geometry can be utilized in shape analysis. A theorem from differential geometry that is particularly useful is [10],

Theorem: Every regular curve $c:I_s \rightarrow R^n$ can be parameterized by arc length. In other words, given a regular curve $c:I_s \rightarrow R^n$ there is a change of variables $\theta : I_s \rightarrow I_\theta$ such that $|(c.\theta)'(s)|=1$, where $(c.\theta)$ is a composite function.

Let $c = c(s)$ be the parametric representation of the curve under analysis with s as the natural parameter (i.e $|dc/ds| = 1$) then the vectors t, n, b, satisfy the Serret-Frenet equations [7], [10],

$$\begin{bmatrix} \dot{t} \\ \dot{n} \\ \dot{b} \end{bmatrix} = \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{bmatrix} \begin{bmatrix} t \\ n \\ b \end{bmatrix} \qquad (2.11)$$

where,

κ = curvature,

τ = torsion,

t = tangent,

n = normal at the point,

b = binormal at the point,

and the dot notation denotes the deravative with respect to the natural parameter s.

When the curve under analysis lies in a plane, the torsion τ in general is equal to zero and the binormal b is constant. Thus for a curve in the x-y plane (2.11) reduces to,

$$\begin{bmatrix} \dot{t} \\ \dot{n} \end{bmatrix} = \begin{bmatrix} 0 & \kappa \\ -\kappa & 0 \end{bmatrix} \begin{bmatrix} t \\ n \end{bmatrix} \tag{2.12}$$

Now if θ is the angle made with respect to the x-axis by the tangent to the curve then,

$$\begin{bmatrix} t \\ n \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} i \\ j \end{bmatrix} \tag{2.13}$$

where i and j are unit vectors in the x and y directions respectively, Differentiating (2.13) with respect to the natural parameter s yields,

$$\begin{bmatrix} \dot{t} \\ \dot{n} \end{bmatrix} = \dot{\theta} \begin{bmatrix} -\sin\theta & \cos\theta \\ \cos\theta & -\sin\theta \end{bmatrix} \begin{bmatrix} i \\ j \end{bmatrix} \tag{2.14}$$

Substituting equation (2.13) in equation (2.14) yields the result,

$$\begin{bmatrix} \dot{t} \\ \dot{n} \end{bmatrix} = \begin{bmatrix} 0 & \dot{\theta} \\ -\dot{\theta} & 0 \end{bmatrix} \begin{bmatrix} t \\ n \end{bmatrix} \tag{2.15}$$

Comparing (2.12) with (2.16)

$$\kappa = \dot{\theta} \tag{2.16}$$

In view of the different representations of the word curvature [2], [21] it is necessary to emphasize that the derivative in (2.16) is with respect to the natural parameter s and not with respect to some arbitrary distance measure. If a curve is not represented in terms of the natural parameter, but some other real valued function (say $\theta = \theta(s)$) then this transformation should be allowable. The implication is that, $\theta : I_s \rightarrow I_\theta$ is an injective mapping of the interval I_s onto I_θ. Where I_s and I_θ are the respective domains in s and θ over which the curve is defined.

Thus the function arctan(θ) with values between $\pm$ 90 or arccos(θ) with θ between 0 and 180 are not allowable changes of parameter and any property based on these transformations may not be a property of the curve but a property of the representation. Thus if a curve or a shape has been represented in terms of the sample number 'k' and if the algorithm is unable to affect the one-to-one transformation described above, then the following fact should be exploited [10],

$$\frac{ds}{dk} = \left| \frac{dc}{dk} \right| \tag{2.17}$$

Using (2.17) it can be shown [10] that the magnitude of curvature can be obtained from the following relationship,

$$|\kappa| = \begin{cases} |c' \times c''| \; / \; |c'|^3 & \text{if } c' \neq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.18}$$

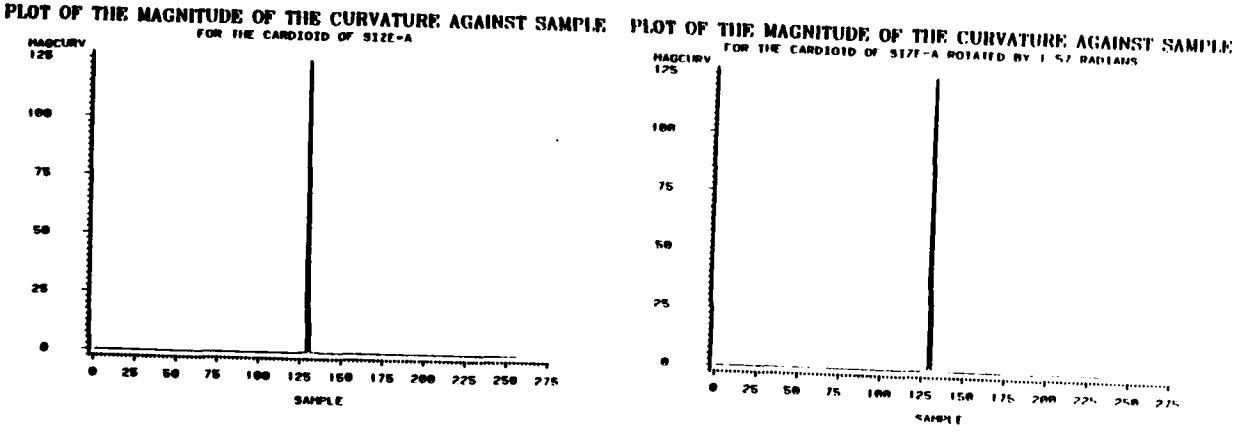


FIG. 6. CURVATURE OF THE CARDIOID SHAPE.

Where the symbol x denotes the vector cross-product and the symbol ' denotes differentiation with respect to k. Curvature is a vector quantity and it points in the direction of the normal.

Using equation (2.18) on the cardioid and its rotated version (Fig. 4a and 4b) resulted in the plots of the curvature magnitudes shown in Fig. 6a and 6b. As predicted the magnitude of the curvature is identical before and after rotation. This is in sharp contrast to substantial differences demonstrated in Fig. 5a and b.

SECTION III
BASIC SHAPE CONCEPTS

In this section several general concepts from allometric disciplines [11] are combined with some entirely new concepts concerning shapes. This combination provides the foundation for a new approach to defining shape vectors in the appropriate shape space. This new space is a combination of properties of vector spaces and is called definitively shape space.

Applying the shape space concepts to the shape analysis problem provides a basis for the recognition that the features of two or more shapes under analysis are the same. For instance, with shape space concepts it is possible to deter-

mine that a partial shape, independent of size or rotation, belongs to a more complex whole shape.

Typically in a shape analysis problem, an algorithm operates on the shape data according to a set of criteria for the purpose of reaching a decision of some sort. Usually the decision is whether or not two or more shapes are the same. The shape analysis algorithm utilizes measurements such as curvature, and measurements between predefined points on the shape, such as length width, diameter, area, etc. Therefore if K shapes are under analysis, and N measurements m_{ki} are made on each shape, then the result is the K measurement vectors,

$$\begin{aligned} M_1 &= (m_{11}, m_{12}, m_{13}, \ldots, m_{1N}) \\ &\vdots \\ M_k &= (m_{k1}, m_{k2}, m_{k3}, \ldots, m_{kN}) \\ &\vdots \\ M_K &= (m_{K1}, m_{K2}, m_{K3}, \ldots, m_{KN}) \end{aligned} \qquad (3.1)$$

where the first subscript of m_{ki} refers to the object being measured while the second subscript i refers to the i th measurement between the predefined points on the k th shape. Any two objects will then be said to have the same shape with respect to these measurements if one vector is a scalar multiple of the other,

$$M_k = a\, M_p \qquad (3.2)$$

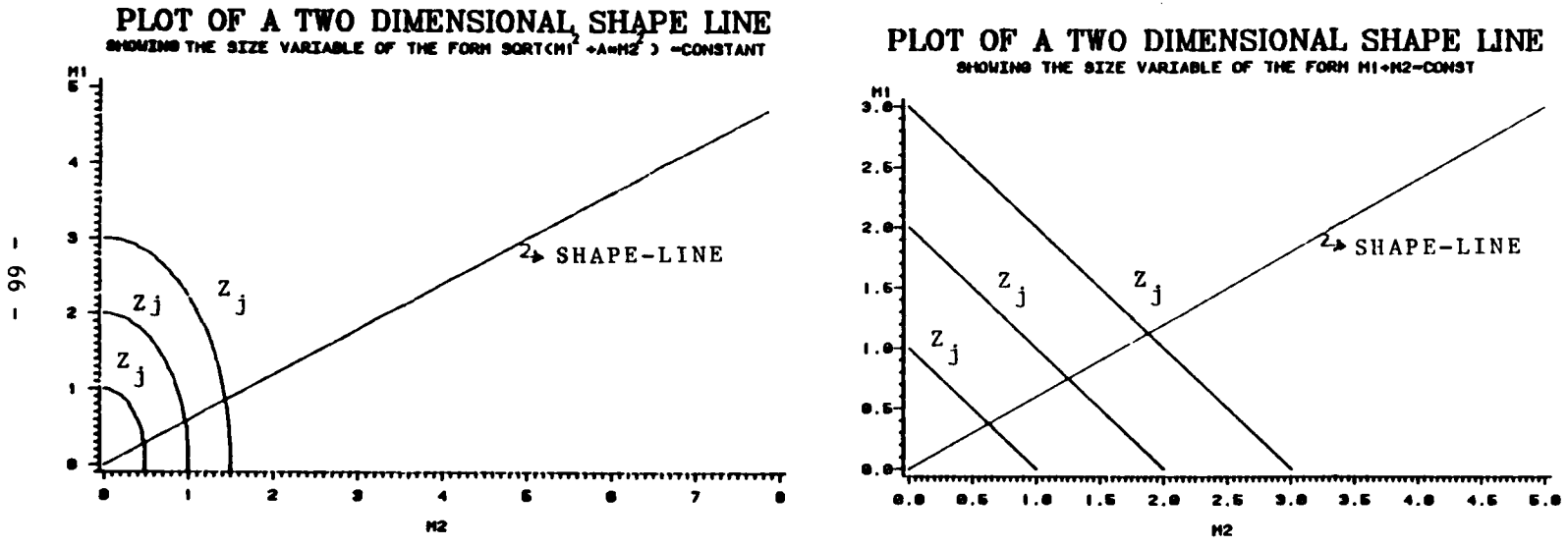


FIG. 7. EXAMPLES OF SIZE-VARIABLE IN 2-D SHAPE SPACE.

where a is a scalar greater than zero.

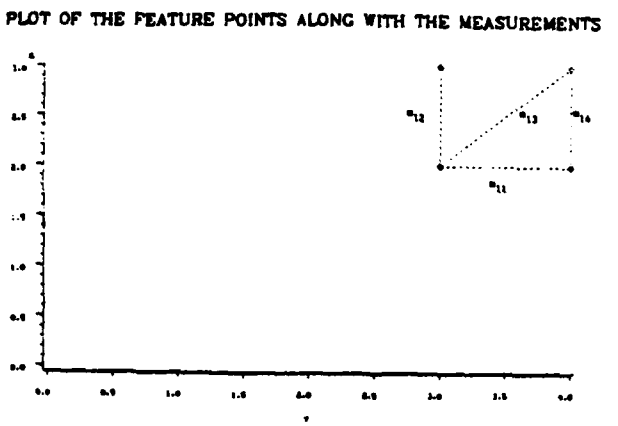
The geometric significance of (3.2) is that in the N dimensional space of positive measurements all points on a straight line through the origin define the same shape. Points can be uniquely located on the positively directed line by finding the intersection of first order surfaces with the line defining the shape. The class of functions which define these surfaces are linear homegenous functions of order one, of the measurements, m i=1, 2,.N. The mathematical representating for this class is,

$$\Xi\ (a\ m_{ki}) = a\ \Xi\ (m_{ki})$$
$$\Xi\ (m_{ki}) \geq 0 \qquad (3.3)$$

where a is scalar

and Ξ refers to a countably infinite class of linear functions, with members $Z_j\ (\ m_{ki}\ )$.

The distance from the origin to the intersection of a particular member of the class $Z_j (\ m_{ki}\ )$ with the ray through origin defining the shape is refered to as the size, scale factor or the normalization factor with respect to that member. Following Mosaiman [11] terminilogy, in the sequel $Z_j (\ m_{ki}\ )$ will be refered to as the size variable. Some examples of size variable in a measurement space of two vectors are shown in Fig. 7a and b.



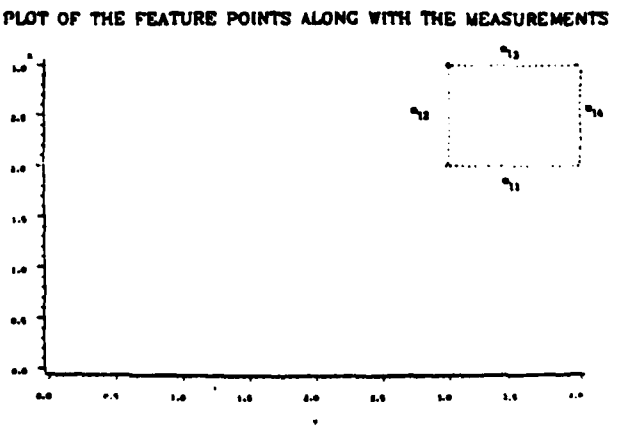


FIG. 8. MEASUREMENTS ON A SQUARE-SHAPE.

These statements merit further discussion and clarification because of their implications. Consider for example the two different measurement vectors extracted from the simple shape shown in Fig. 8. The shape is a unit square. In the first case, the measurement vector is,

$$M_1=(m_{11} , m_{12} , m_{13} , m_{14})=(1 ,1 ,1.414, 1) \quad (3.5)$$

while in the second case, the measurement vector is

$$M_2'=(m'_{21} ,m'_{22} , m'_{23} , m'_{24})=(1, 1, 1, 1) \quad (3.6)$$

where the ' in the above equation indicates that the measurements are between a different set of feature points of the shape. Now if,

$$Z_j^i (m_{ki}) = m_{k3} \quad (3.7)$$

is chosen as the size variable then the corresponding shape vectors are,

$$S_1 =(.707 , .707 , 1 , .707) \quad (3.8)$$

$$S_2 =(1 , 1 , 1 , 1)$$

Comparing these two shapes without any reference to the size variable or the points between which the measurements were made, one would conclude that the the two shapes are not the same. Thus not only has the functional form of the size variable to be the same but the measurements involved in the

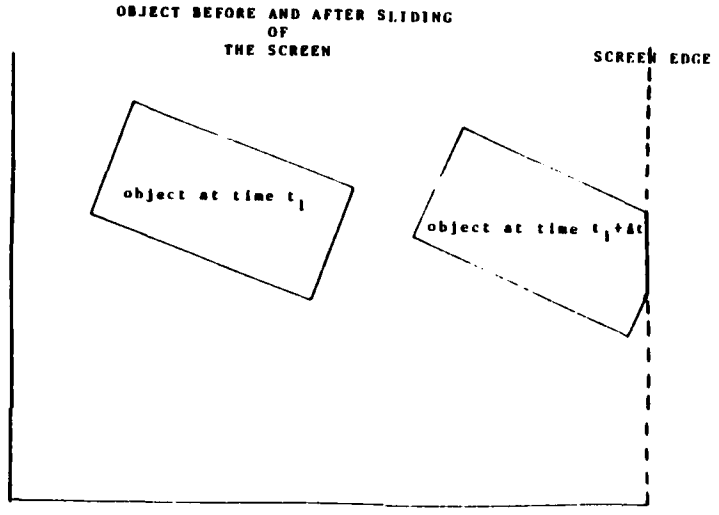


FIG. 9. OBJECT LEAVING THE FIELD OF VIEW.

functional relationship have to be between the same feature points. Now consider the shapes in Fig. 9. This is the typical situation in which occlusion occurs or the shape is outside the field of view of a camera (or some other measuring device). Assume that both shape boundaries are represented by an equal number of samples and that every point on each shape boundary is defined as a feature point. Now if the standard deviation of the data is chosen as the size variable then it is obvious that the standard deviation of the two shape boundaries are different. Therefore, this is a case where the size variable is dependent on the shape, which implies that the shape vector is dependent on the size variable. Comparision of the two shape vectors with such a size variable is bound to lead to errors. These problems can be alleviated by defining the quantities in the proper context. This can be accomplished only if the variables are defined in the proper space.

SHAPE SPACE

In this section, the concept of a shape space is introduced. The space is defined in terms of its properties in the usual manner and then two theorems addressing the problem of partial shapes are stated and proved. Assuming that $S_{kj}(s_{ki})$ is a shape vector consisting of well defined operations on the measurements of the shape under consideration.

It will be shown in the sequel that it is necessary to choose a size variable that is independent of the shape under comparison. A necessary first step is the definition of shape vector [11].

SHAPE VECTOR

The Shape vector S_{kj} is defined to be the ratio of the measurement vector M_k to the size variable Z_j.

$$S_{kj} = M_k / Z_j$$

$$S_{kj} = \left(\frac{m_{k1}}{Z_j(m_{ki})}, \frac{m_{k2}}{Z_j(m_{ki})}, \frac{m_{k3}}{Z_j(m_{ki})}, \ldots, \frac{m_{kN}}{Z_j(m_{ki})} \right) \qquad (3.4)$$

It may be noted that the first subscript of S_{kj} corresponds to the object whose shape is under consideration while the second subscript corresponds to the size variable chosen from the class.

Here the points to be emphasized are that,

1) All measurements are made between pre-defined points.

2) A shape vector is defined with respect to a size variable Z_j (m_{ki}). Thus only shape vectors defined with respect to the same size variable can be compared.

3) If two shape vectors are equal, then the two objects have the same shape with respect to the measurements.

4) The shape vector should be independent of the size variable.

These measurements are the m_{ki}'s previously defined; the elements of $s_{kj}(s_{ki})$ are obtained by the following operation,

$$s_{ki} = m_{ki} / Z_j (m_{ki}) \quad (3.9)$$

The shape vector $s_{kj}(s_{ki})$ must satisfy the following properties in addition to the properties of normal Euclidian space.

PROPERTIES OF SHAPE SPACE

1) The shape vector is independent of the size variable.

$$S_{kj} (a s_{ki}) = S_{kj} (s_{ki}) \quad (3.10)$$

where i = 1, 2,.....,N, k=1, 2, ... K,

and a is a scalar.

2) The shape vector is independent of translation that is,

$$S(s_{kj} + s_{ki} + s_0) = S(s_{kj} s_{ki}) \quad (3.11)$$

where i = 1, 2 , ...N, k=1, 2, ...K,

and s_0 is a constant vector.

3)The shape is independent of rotation.

$$S(\alpha_{kj} + u_0) = S(\alpha_{kj}) \quad (3.12)$$

where i=1, 2, ...N and α_i is the angle made by the i th component of S with a fixed reference axis and u_0 is a constant angle.

The vector obtained by using a set of measurements on a partial shape must still be contained in the space.Unless a size variable is found which is independent of both shapes (both the partial and the complete) it is not meaningfull

to compare the shape vector in shape space. It is not possible to find a size variable which is a totally independent continious function of measurements made on both shapes. This being the case the only choice left is to split the shape into parts or subshapes and define a size variable which is piecewise continuous over these parts. Two theorems relating the subshape to the complete shape in shape space are stated and proved. These theorems are used extensively in the sequel.

THEOREM 1 : The vector formed by concatenating a series of shape vector is a shape vector.

Proof: It is required to show that the shape vector resulting from concatenating a series of shape vectors satisfy the three properties of shape space.Let the concatenated shape vector be

$$S_{cj} = (S_{1k}\ (\ s_{1u}\ ),\ S_{2\ell}\ (\ s_{2v}\ ),\ \ldots,\ S_{on}\ (\ s_{ow}\ ),\ \ldots) \tag{3.13}$$

where the first subscript on S represents the object(shape) and the second represents the size variable used to obtain that shape vector. It should be pointed out that k, ℓ and n are arbitrary size variables. Each element of S_{cj} (s_{ci}) can be represented by

$$s_{ci} = m_{oi} / z_d\ (\ m_{ci}\ ) \tag{3.14}$$

Therefore,

$$S_{cj}(s_{ci}) = (s_{ci}, s_{c2}, s_{c3}, \ldots\ldots, s_{cn}, \ldots.) \tag{3.15}$$

Multiplying each component of (3.13) by the scalar a yields,

$$a\,S_{cj}(s_{ci}) = (a\,S_{1k}(s_{1u}), a\,S_{2l}(s_{2v}), \ldots, a\,S_{3m}(s_{3w}) \ldots) \tag{3.16}$$

but from the (3.10) we have,

$$S_{ov}(a\,s_{op}) = a\,S_{ov}(s_{op}) \tag{3.17}$$

using (3.17) in (3.15) results in

$$a\,S_{cj}(s_{ci}) = (a\,s_{c1}, a\,s_{c2}, \ldots, a\,s_{cn}, \ldots) \tag{3.18}$$

or

$$a\,S_{cj}(s_{ci}) = S_{cj}(a\,s_{ci}). \tag{3.19}$$

Equation (3.19) proves property i)

To prove that the shape vector satisfies proverty ii) it is only necessary to observe that each member of the concatenated vector is a shape vector is a shape vector.Therefore each satisfies:

$$S_{kj}(s_{ki} + s_0) = S_{kj}(s_{ki}) \tag{3.20}$$

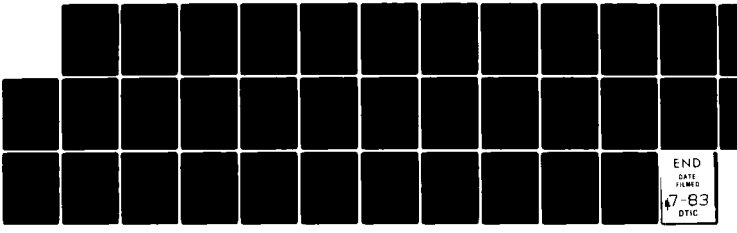
The proof for property iii) follows in a similar manner

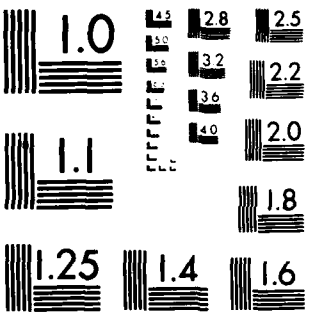
Q.E.D

Theorem2: The shape defined by a shape vector obtained by concatenating a series of shape vectors is unique if and only if the size variable defining each member of the set is known.

AD-A129 221 ADAPTIVE HYBRID PICTURE CODING(U) ARKANSAS UNIV FAYETTEVILLE DEPT OF ELECTRICAL ENGINEERING R A JONES ET AL. 05 FEB 83 AFOSR-TR-83-0499 2/2

UNCLASSIFIED AFOSR-77-3456 F/G 12/1 NL





Proof: → A linear homogeneous function surface will intersect a positive directed straight line at only one point. Since a shape is defined as a point on the shape ray, if follows that any point on this ray can be uniquely determined by its intersection of a size variable which is defined as a linear homegeneous function over the positive quadrant.

: ← Assume otherwise. Then there is at least one S_{on} (s_{ow}) shape vector in the concatenated set whose size variable $z_j(m_{ki})$ can be choosen arbitarily. Then the concatenated vector under this assumption would still satisfy the properties of the shape space, i.e.,

$$a\ S_{cj}\ (\ s_{ci}) \ = \ S_{cj}\ (\ a\ \ s_{ci}) \qquad (3.21)$$

Now the R.H.S of (3.21) can also be expanded as

$$= (\ a\ S_{1k}\ ,\ a\ \ S_{2p}\ ,\ \ldots \qquad ,a\ S_{on}\ \ldots.) \qquad (3.22)$$

$$= \left(\ \frac{aM_1}{Z_k} \qquad ,\frac{aM_2}{Z_p} \qquad ,\ldots. \qquad .,\frac{a\ M_0}{Z_n}\ \ldots\ \right) \qquad (3.23)$$

but since Z can be choosen arbitrarily as long as it satisfies the definition of a size variable(3.4). Choose

$$Z_h = \ Z_h\ (\ a\ m_{oi}\ ) \qquad (3.24)$$

substituting (3.4) and (3.24) if (3.23) the following relationship is obtained

$$a\ S_{cj}(\ s_{ci}\ ) = \left(\ \frac{a\ M_1}{Z_k} \quad , \frac{a\ M_2}{Z_p} \quad \ldots \quad , \frac{M_0}{Z_n}\ \ldots\ \right) \qquad (3.25)$$

which is also equal to,

$$= \quad (a\,S_{1k}\,,\ a\ \ S_{2\ell}\,, \qquad \ldots\ S_{on}, a\,S_{0+1,\,n+1}\ldots) \qquad (3.26)$$

comparing (3.22) with (3.26) that,

$a\,S_{cj}(\,s_{ci}) \neq S_{cj}\,(\,a\ \ s_{ci})$

which contradicts (3.21)

Q.E.D

SECTION IV

CRITICAL POINTS

When analyzing shapes represented by sampled boundries a natural question is one regarding the points between which the measurements should be made. A simplistic approach, an approach which assumes that there are the same number N of boundary points on each shape and evades the problem of analyzing shapes with a different number of boundary points, would be to make measurements between all possible pairs of boundary points. This leads to the number $N!/(N-2)!2!$ of measurements for each shape.

In order to recognize the shape, a normalization and comparison procedure must still be used. It is obvious that for relatively small N the computational requirements are very large. Therefore it becomes necessary to consider measurements between a relatively few representative points on the shape. The chosen points on the shape have to be more impor-

tant, in some sense, than all the rest. The set of points which define a shape may be considered as a Fuzzy set in which various points are assigned to it with various degrees of membership. Defining a precise criteria on which a degree of membership can be assigned to a point on a shape is very difficult. However, past researchers have discovered [2], [8], [18] that, the points that should have a higher degree of membership than the rest are:

1) Points of Maxima
2) Points of Minima
3) Points of Inflection
4) Points of Intersection
5) End points of open curves
6) Points where the curvature changes sign or magnitude.

In past shape analysis efforts, points of maxima as well as minima, and inflection points of curves have been extracted from the shape data without due consideration of the coordinate axes. This approach inevitably lead to errors, because these points have no meaning unless the coordinate axes are first defined. The problem with such an approach is that if a set of coordinate axes is chosen independent of the shape under consideration then the maxima, minima, and inflection points will not be independent of rotation of the shape with respect to the coordinate axes. The conclusion

is that the coordinate axis upon which the maxima, minima, and points of inflection are based must be dependent on the shape itself.

The next logical question is whether there should be one set of coordinate axes or many. The answer is not straightforward. Some shapes require more than one set of coordinate axes while others may require only one. Before discussing methods that may be used for determining critical points, two definitions are presented which will be used in the sequel.

Definintion I: A curve c is said to be in line of sight of a point P if every point on the curve c can be connected without intersecting the curve at any other point. Otherwise the curve is said to be not in line of sight of the point P (NLSP). Examples of (LSP) curves of a point P, where P is the centroid of the shape then denoted as (LSC), are shown in Fig. 1Ø.

Definition II: Line of sight of a straight line axis L: Let n be the normal projection of a curve c onto a straight line L. The curve c is then said to be in line of sight of L if all points from c can be can be mapped injectively onto n.

Examples of curves which are in line of sight of a single straight line axis (LSA) are shown in Fig. 11 along with

EXAMPLES OF SINGLY CLOSED CURVES

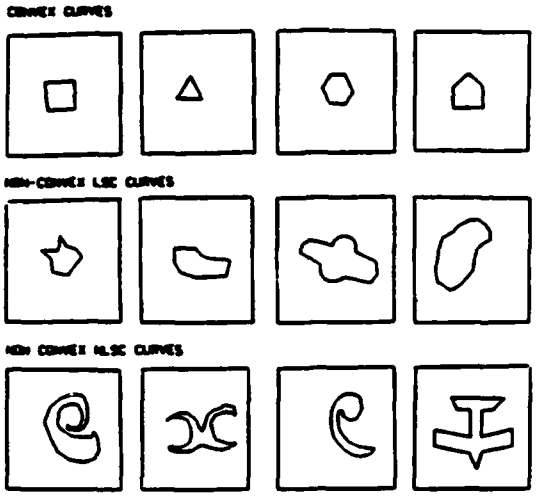


FIG. 10. CURVES DEPICTING THE LINE OF SIGHT OF A POINT CONCEPT.

EXAMPLES OF SINGLY OPEN CURVES

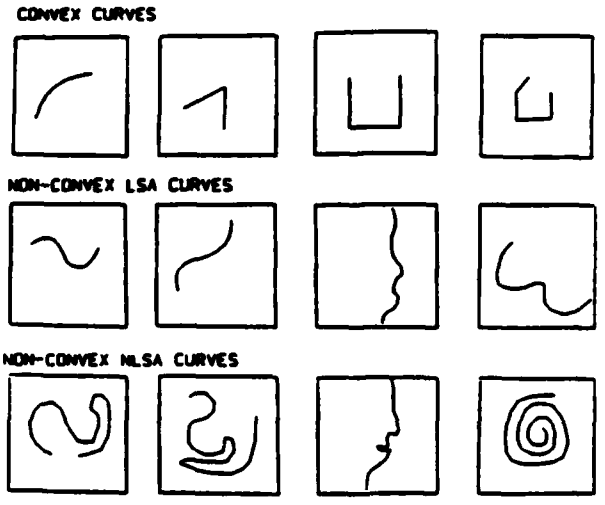


FIG. 11. CURVES DEPICTING THE LINE OF SIGHT OF AN AXIS CONCEPT.

curves not in line of sight of a single straight line axis(NLSA)

Dividing the shape into this set of projections is equivalent to defining the shape in terms of single valued functions. It follows from the definition of a single valued function that fewer ambiguities should result. That is, the critical points are now determined from single valued functions rather than multivalued functions. The actual methods for determining the critical points can now be presented.

SECTION V

ADAPTIVE LINE OF SIGHT CRITICAL POINT DETERMINATION

Several methods have been utilized in the past for determining critical points [2],[8],[18]. Most of these methods are based on operations on a fixed number, m, of adjacent points. Since it is unlikely that an intelligent machine will have a priori knowledge about the size of the shape to be analyzed or the relationship of the number of sample points to the size, the ambuiguities involved in detection of critical points using methods that are totally dependent on operations on fixed number of adjacent points, are high. The classical methods of differential geometry work very well on one dimensional data and on theoretical curves. While each of these methods has exhibited desirable charac-

teristics, undesirable characteristics are always present, specially when the shape is corrupted by noise. A brief description of some of the well known classical methods [7],[10], that work well for theoretical curves are now presented for the purpose of comparing their performace with the approach described in the sequel, namely, the Adaptive Line of Sight Method. Also these methods may be used in the second pass of the this adaptive algorithm after the shape has been rid of noise.

The method of Centroidal Vectors

In this method the ith point on a shape at a vector distance $d_i \angle \alpha_i$ from a reference point (typically the shape centroid) is said to be critical with respect to the reference point if

$$((\,|d_{i-i}|-|d_i|\,) \text{ and } (\,|d_i| - |d_{i+1}|))$$

or

$$((|\alpha_{i-1}|-|\alpha_i|\,) \text{ and } (\,|\alpha_i| - |\alpha_{i+1}|\,))$$

have opposite sign.

This operation is very local in nature and is extremely sensitive to noise. Round-off or truncation errors also have a deleterious effect on the operation. It also fails very often when dealing with smooth curves or in situations where the centroid is located away from the shape boundary as shown in Fig. 12.

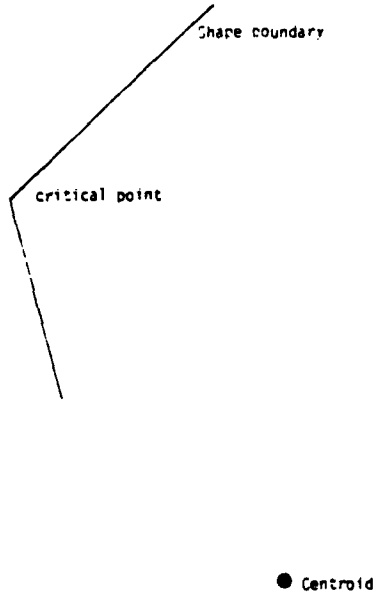


FIG. 12. A CASE AGAINST THE CENTROIDAL METHOD.

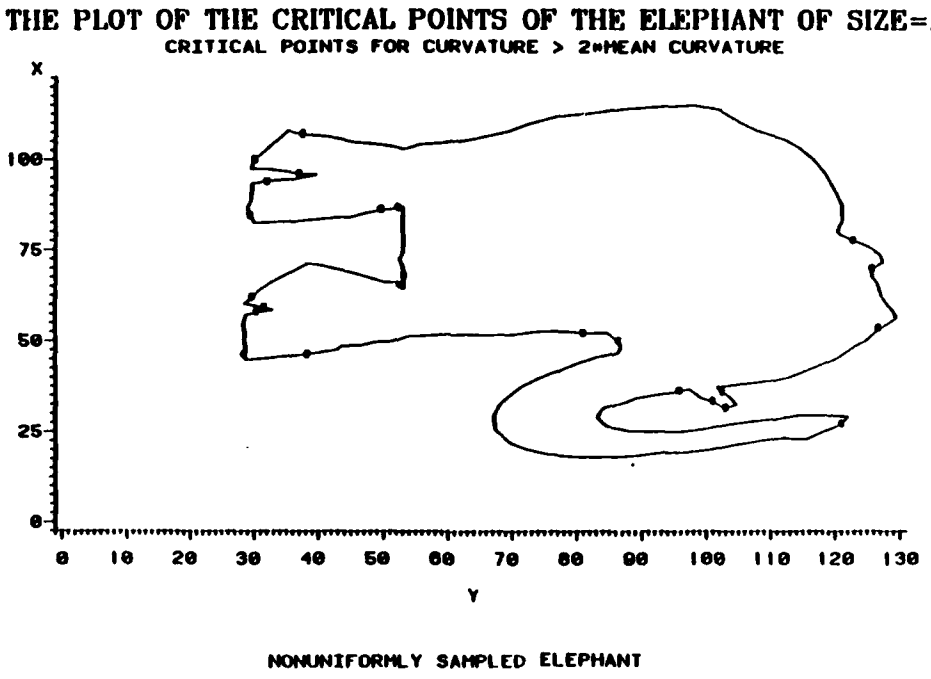


FIG. 13. CRITICAL POINTS FOR THE ELEPHANT-SHAPE.

Another method is that of Curvature Vectors. This method is basically a two pass process. In the first pass, the i th point on the shape with a curvature κ_i is set as a critical point if

$(\kappa_{i-1} - \kappa_i)$ or $(\kappa_i - \kappa_{i+1})$

have opposite sign. In the second pass, the points at a maximum distance from the straight line joining every two successive critical points is set as critical. This method like the previous method, is very sensitive to noise. Though it works for most of the smooth curves, there are instances where it fails. For example, this method produces only one critical point for the cardioid discussed in section II.

In the Magnitude of Curvature Method the ith point is determined to be critical if,

$|\kappa_i| > =$ Threshold

The critical points were determined by using this method on the shape (elephant) in Fig. 13. It is apparent that some critical points were missed. For example the critical points for the trunk are missing. This procedure is more immune to the noise than the others; however an equally critical problem is added. Specifically, a threshold must be determined a priori. Additionally the method fails completely on smoothly varying curves such as the cardioid of Fig. 5.

A procedure that performs well for polygonal shapes and curves with very few boundary points is called the Line of Sight method. In this method, if d' denotes the normal distance of the i th point from a straight line L, then the i th point on the shape boundary is said to be critical with respect to the straight line L if

$(d_{i-1} - d_i)$ and $(d_i - d_{i+1})$

have opposite sign. The set of critical points found with respect to the set of tangent lines drawn at all points on the shape, will then be called the critical points. It is obvious that for shapes other than polygonal an infinite number of axes or tangent lines are required. Nonetheless, the attributes exhibited by this method are very desirable. It is however necessary to reduce the dimensionality of the problem; this is the subject that is addressed by the Adaptive Line of Sight method.

ADAPTIVE LINE OF SIGHT METHOD:

In the Adaptive Line of Sight method the critical point determination is based on a set of coordinate axes that are dependent on the shape under consideration. As previously discussed, this will allow critical points to be determined with fewer ambiguities. The Adaptive Line of Sight method is an approach used to emulate the Line of Sight method described above without having the burden of

excessive dimensionality. The procedure is adaptive in the sense that it adapts to the shape data under consideration.This will become evident in the sequel.

The Adaptive Line of Sight method is a two pass process. In the first pass, the shape is divided into a minimum number of segments or parts by an appropriate set of 'critical points'. The members of this segmenting set of critical points are defined to be those critical points, such that all boundary points in between any two adjacent critical points have the following two properties with respect to the straight line L joining critical points:

1)the boundary points are on the same side of a straight line L joining the adjacent critical points;

2)the points are in line of sight of L.

From the definition of line of sight (Definition II) it is clear that this means that the boundary curve has a unique one-to-one projection on L. It should be emphasized that the minimum number of critical points are obtained during the first pass. This is done by an exhaustive search process that locates all points such that the above properties are satisfied. Often the minimum is not unique, in which case, the required set is obtained by summing all minimum sets.

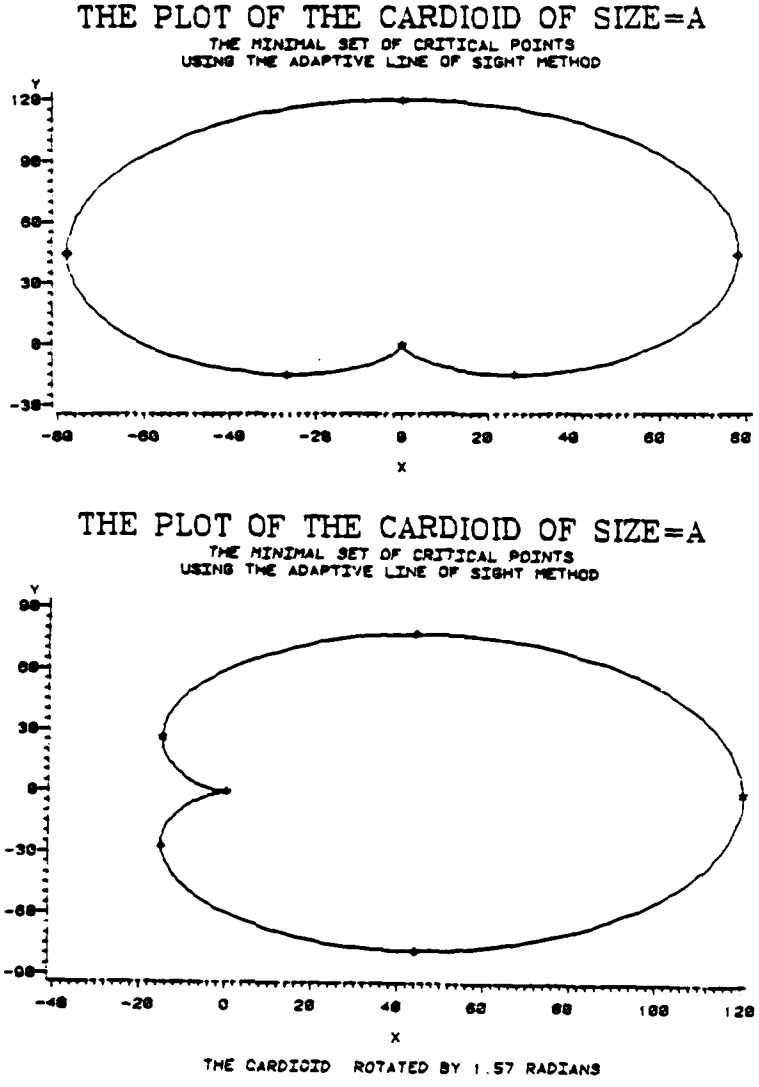


FIG. 14. CRITICAL POINTS OF THE CARDIOID BY THE ADAPTIVE METHOD.

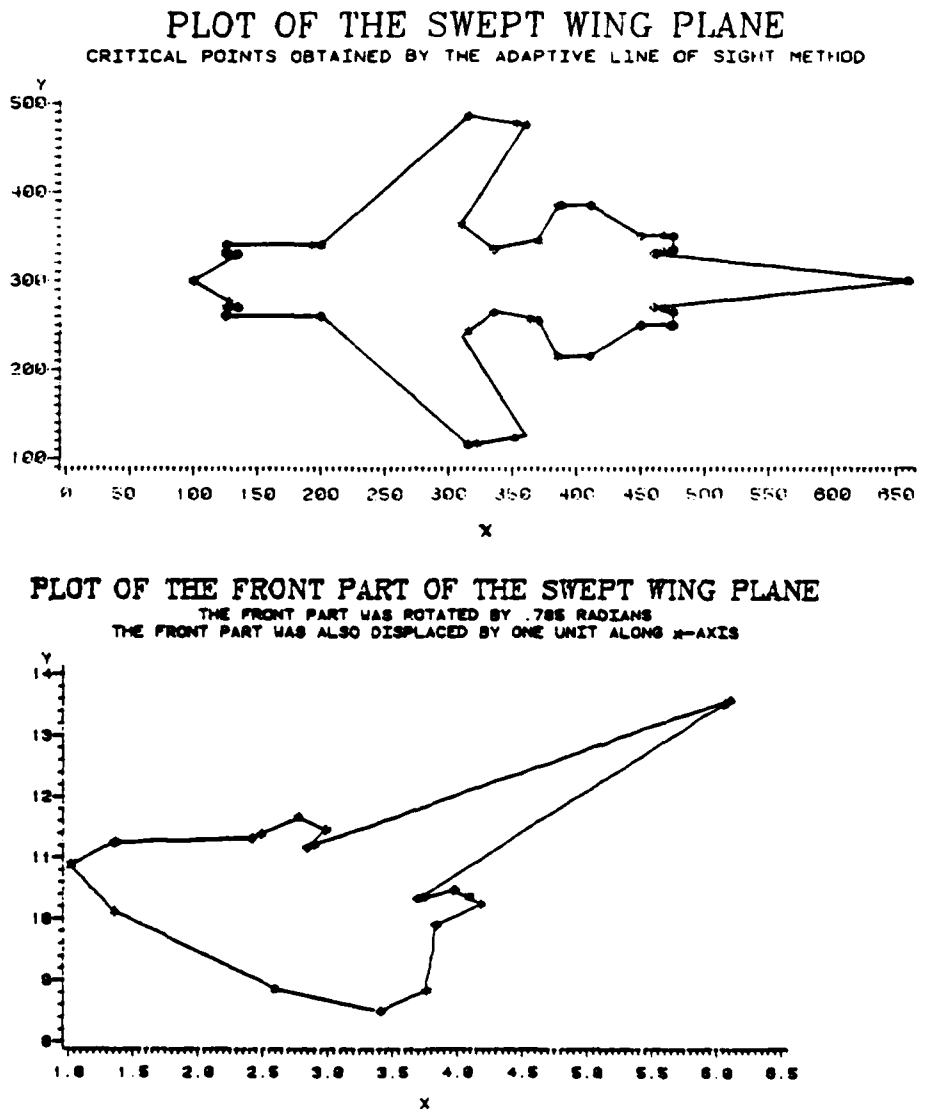


FIG. 15. CRITICAL POINTS OF SWEPT-WING PLANE SHAPES.

In the second pass, the points of maxima - minima and inflection between every two critical points are detected using the derivative of the normal distance of the point from L with respect to the distance along L. A moving average of these normal distance may be used to eliminate the effect of noise. The critical points found in pass one, the segmenting set, and the critical points found in pass two are defined to be the members of the Fuzzy shape set with the highest degree of membership. A thorough description of both the computational and detection aspects of the Adaptive Line of Sight Algorithm along with a complete flowchart, is given in Appendix A.

The minimal set of critical points obtained by using the Adaptive Line of Sight on the cardioids of Fig. 4, as predicted are independent of rotation and are shown in Fig. 14. Fig. 15 shows a typical non minimal set of critical points obtained using the Adaptive Line of Sight method. Part a) of the figure is the complete shape of the plane, while part b) shows the front part of the plane after it has been rotated, displaced and reduced in size. In spite of the total dissimilarity, the critical points found for both the partial plane shape and the whole plane shape are quite close as predicted. The clusters of critial points occur because of the lack of a predefined resolution.

SECTION V I

FEATURE SELECTION & COGNITIVE STEP

The features of a shape are essential to defining the shape in terms of parameters that can be ultimately used by machine for decision purposes. However, the manner in which the feature defining procedure can be selected is quite variable. Since a dependable feature selection procedure is fundamental to the shape recognition problem, it is essential that this aspect of shape recognition be addressed with specificity. This point is punctuated when it is realized that, irrespective of the method of defining and detecting critical points, a cognitive algorithm is still required which examines in some sense, the critical points of the shape for the purpose of reaching a decision about some aspect of the shape. Consider, for example, the shape shown in Fig. 15-a, a shape such as this swept wing plane may have thirty to fifty critical points. The human eye makes numerous measurements, automatically and sub-consciously, between the feature points and determines their relationship to one another. The 'most important' of these measurements combined with the relationship between them, comprise the decision set. The term " most important " is difficult to define mathematically, because it is the

result of training. An unrefined cognitive procedure must therefore consider the set of all possible measurements between the critical points. Obviously this is a very large number of measurements even for numbers as modest as thirty to fifty. It is well known that observation that this totality of measurements between critical points is not essential to the decision process.

It is necessary therefore to determine methods for acquiring the minimal set of measurements or features required for the decision process. The human, apparently places heavy weighting on features that are formed by critical points that are symmetrically opposite about an axis and features that are extracted from adjacent critical points concerning the shape. Without any prior knowledge a human can find the sets of axes about which some critical points are symmetrically placed with very little effort. However, such a task is almost insurmountable for a machine based algorithm unless it is performed at a post-cognitive level. In the absence of noise, machine recognition (cognitive) algorithms perform reasonably well by using only features consisting of adjacent critical points.

A cognitive algorithm that utilizes measurements such as these in a continuous sequential manner would be en-

tirely adequate if the algorithm for detecting critical points is totally immune to noise, round off, and truncation errors. For example any extra critical points that are the result of a burst of noise would prevent any continuous sequential recognition algorithm from yielding conclusive results.

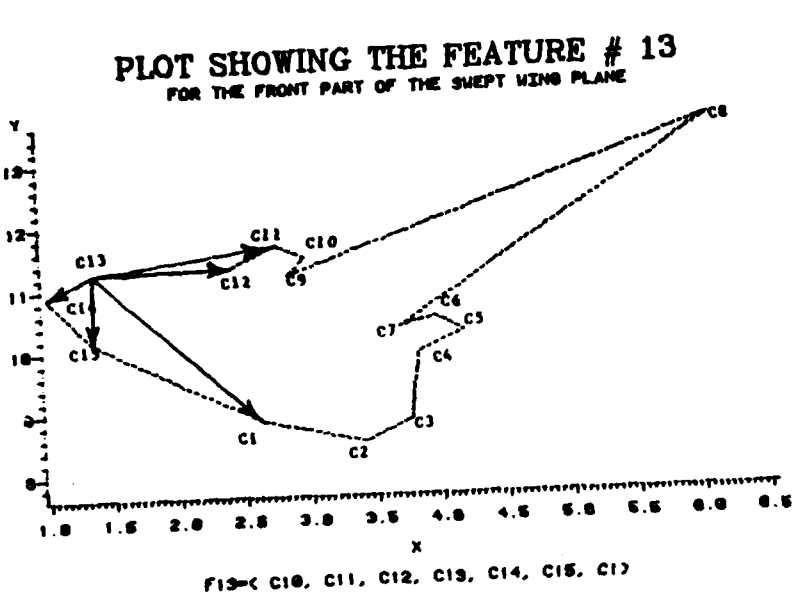
One manner by which this problem can be circumvented is to divide each of the shapes under analysis into subshapes in terms of their features, and then compare the features of these subshapes and the manner in which they are related to each other. It is recalled from Section III that the properties of shape space dictate that the measurements which define a feature must be made from either the centroid of the set of critical points or between the critical points that form the feature. It is necessary then to determine, in some manner, the minimum number of critical points that adequately define a feature. If each feature were defined in terms of only two critical points, then all features would be identical since the comparisons are made with respect to the same shape independent size variable. Therefore, the minimum number of critical points that can form a distinguishable feature is three, and these must be adjacent. However unless the relationship between these three point features with its adjacent fea-

tures is also, considered any comparison (cognitive) algorithm almost always leads to ambiguities. The reason for this is because the three point feature forms a triangle. An examination of the shape shown in Fig. 15-a shows that it contains many similar triangles.

Unfortunately the mathematics required to obtain the optimal number of critical points that should form a feature is not yet developed. Therefore, it is necessary to resort to the psychological aspects of the human recognition and decision process as well as the practical aspects such as the implementation and computational requirements. These criteria lead to the features being selected as follows:

1) Reconstruct the shape by connecting all the adjacent points by a straight line. This is called the critical shape boundary.

2)The feature F_i is then formed by including critical points, C_i and at most three adjacent critical points on each side of C_i. The critical points chosen must be in line of sight of C_i . This means that it must be possible to draw a straight form C_i to each of the other critical points defining the feature without intersecting the critical shape boundary. The feature obtained by this procedure cannot be defended as optimal in any mathematical



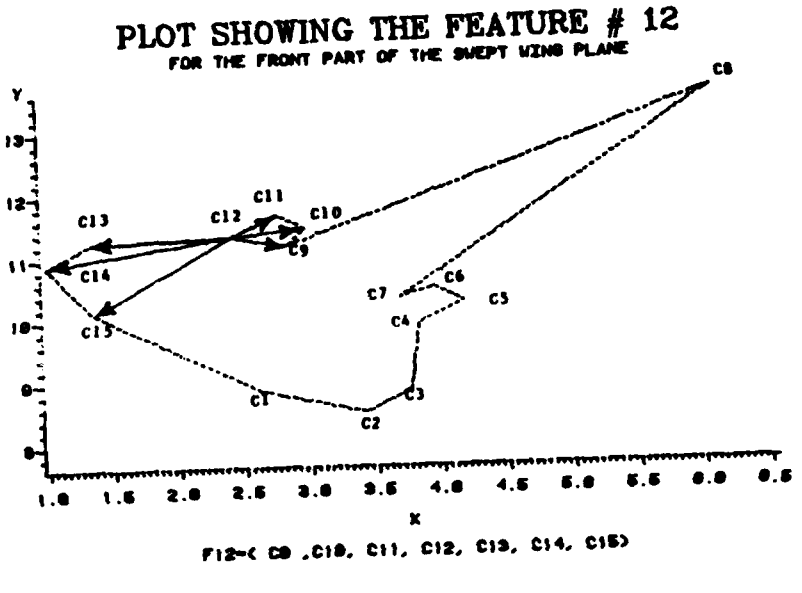
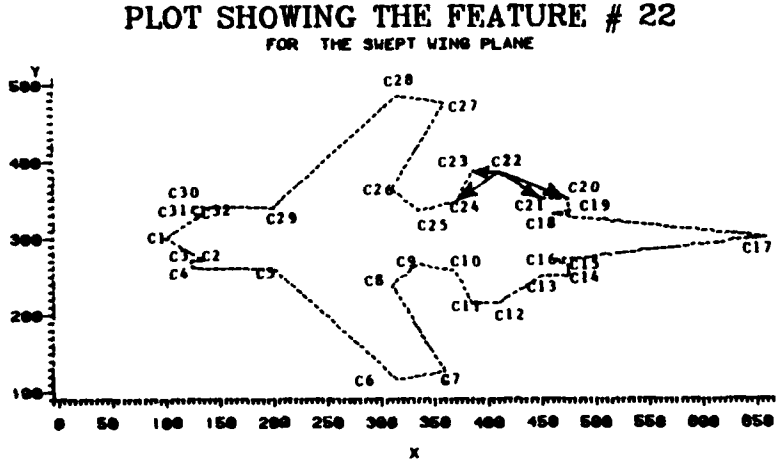
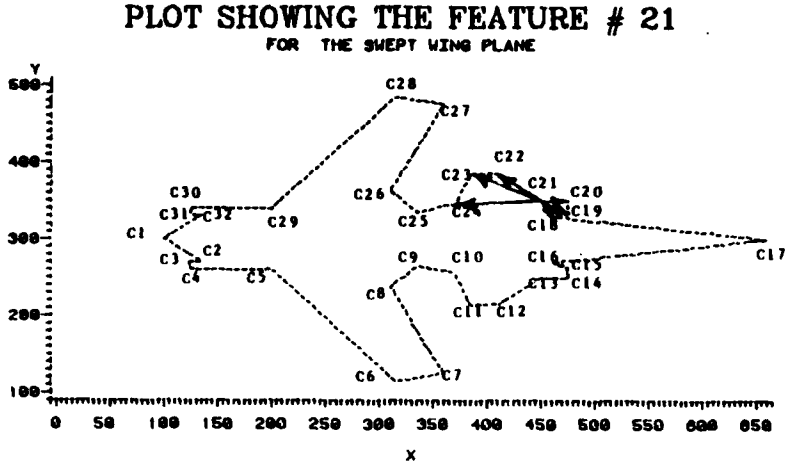


FIG. 16. CRITICAL POINTS AND EXAMPLES OF FEATURES FOR THE FRONT PART.



F22=(C28, C21, C22, C23, C24)



F21=(C18, C19, C20, C21, C22, C23, C24)

FIG. 17. CRITICAL POINTS AND EXAMPLES OF FEATURES FOR THE PLANE.

sense. However, it correlates quite well with those entities that humans consider features. The features corresponding to critical points C_{12} and C_{13} are shown in Fig. 16.

A desirable improvement to the above feature defining procedure is an algorithm for deciding whether the line joining C_i to another critical point in the feature lies inside or outside the shape.

The cognitive step requires, as usual a dictionary of the features of the complete shape against which the partial shapes are to be compared. The partial shape dictionary will henceforth be referred to as the problem text. One page of the complete shape dictionary is shown in Table 1. This page contains features twenty-one and twenty-two of the swept wing plane of Fig. 17. The table includes, in addition to the feature number, the critical points of that feature along with their x and y location, the x and y location of the centroid of all the critical points contained in the feature, the size of the feature, the normalized components of the shape vector, and the normalized angle of the shape vector component.

The normalized shape vector component are defined $s_{ki} \angle \alpha_{ki}$ where s is obtained by the relationship,

$$s_{ki} = \text{sqrt}\ ((x_{ki}-\bar{x})^2 + (y_{ki}-\bar{y})^2)/\ Z_j \qquad (6.1)$$

where the size variable was chosen to have the form,

$$z_j = \sum_{i=1}^{n} \text{abs}(x_{ki} - \bar{x}) + \text{abs}(y_{ki} - \bar{y}) \qquad (6.2)$$

and the subscript is the feature number. The normalized component of the angle of the shape is obtained in a similar fashion. These same quantities are obtained to form a dictionary for the partial shape (problem text). Each page of both the dictionaries begins with a feature set that is not a subset of another feature set. This feature set is defined as the uncovered feature set. The covered feature sets are arranged in the order of cardinality below the uncovered feature set on each page. The purpose of this architecture is to simplify the computational requirements for the cognitive step.

In general all the features in the dictionary will not be contained in the problem text. It is also true that the problem text contains features that are not present in the dictionary. This becomes apparent by examining Fig. 9. Therefore, the fact that a feature is contained in the problem text, does not imply that the partial shape is not a part of the complete shape, because it is not necessary for the partial shape to have fewer points that the whole shape. Therefore, further examination is required before a decision regarding the problem text can be made.

TABLE 1

A PAGE OF THE DICTIONARY SHOWING FEATURES 21 & 22

CRITICAL POINT	COORDINATE		FEATURE VECTOR	
SEQUENCE #	X-LOCATION	Y-LOCATION	DISTANCE	NORM. ANGLE
C_i	$X_{k,i}$	$Y_{k,i}$	$s_{k,i}$	$\alpha_{k,i}$
18	460	330	0.095814	-.0666
19	475	335	0.121843	-.0746
20	475	350	0.111665	-.0538
21	450	350	0.047611	-.0278
22	410	385	0.098166	-.0650
23	385	385	0.145874	-.0964
24	370	345	0.162900	-.0615

FEATURE # k = 21
SIZE = 385.71 , ANG MEAN = .6889
CENTROID COORDINATES
X = 432.143 , Y = 354.286

20	475	350	0.219780	0.1884
21	450	350	0.129848	0.0269
22	410	385	0.088005	-0.6084
23	385	385	0.149101	-0.1748
24	370	345	0.192721	0.7716

FEATURE # k = 22
SIZE = 266 , ANG MEAN = -.4126
CENTROID COORDINATES
X = 418 ,Y = 363

TABLE 2

A PORTION OF THE PROBLEM TEXT SHOWING FEATURES 12 & 13

CRITICAL POINT	COORDINATE		FEATURE VECTOR	
SEQUENCE #	X-LOCATION	Y-LOCATION	DISTANCE	NORM. ANGLE
C_i	$X_{k,i}$	$X_{k,i}$	$s_{k,i}$	$\alpha_{k,i}$
9	2.8429	11.1716	0.096285	-.0644
10	2.9844	11.4543	0.122441	-.0701
11	2.7724	11.6666	0.112214	-.0485
12	2.4187	11.3131	0.047845	-.0258
13	1.3580	11.2429	0.098648	-.0646
14	1.0043	10.8894	0.146590	-.0928
15	1.3575	10.1115	0.163700	-.0513

FEATURE # k = 12
SIZE = 7.6766 , ANG MEAN = 1.38
CENTROID COORDINATES
X = 2.1054 , Y = 11.1213

11	2.7724	11.6666	0.060616	-0.0535
12	2.4187	11.3131	0.037482	0.0308
13	1.3580	11.2429	0.037192	0.0596
14	1.0043	10.8894	0.043854	0.0878
15	1.3575	10.1115	0.037192	0.0463
1	2.6298	8.8382	0.091575	0.1094

FEATURE # k = 13
SIZE = 21.5099 , ANG MEAN = 0.13964
CENTROID COORDINATES
X = 1.92349 , Y = 10.6769

The decision procedure consists of selecting an arbitrary word from the problem text dictionary. A problem text word is of course a feature from the partial shape under comparison. The shape vector from the problem text is compared to the shape vector in the dictionary always starting on page one of the dictionary. The comparison continues until a match occurs. The next step is to compare the next problem text feature vector in order of , cardinality, to the next feature in the dictionary and so forth. An example of this technique is given by comparing Tables 1 and 2. In this experiment a feature vector from the partial shape was selected for comparison. It should be emphasized that the feature vector is from the problem text of the partial swept wing aircraft shown in Fig. 16. The partial shape has been rotated and shifted as well as scaled to insure that any direct template matching procedure will fail. This also demonstrates that the concatenated feature vector matching procedure described here is independent of rotation, size and location. Feature vector twelve (word) was arbitrarily selected from the problem text. By comparing tables 1 and 2 it is apparent that the word 12 matched feature vector 21 of the dictionary. It should be noted that this match occurs even though the location of the critical points and the centroid of the feature of the partial shape are different from those same

quantities for the whole shape because of the rotation and shift. In this way a correspondence table is then established between the critical points of the features in the dictionary to the critical points of the word in the problem text The next step is to proceed in sequential order to the next word in the problem text e.g. word 13 (n+1) which is sequentially next to word 12 (n) of the problem text does not match feature vector 22 (m+1) of the dictionary. However since feature vector 22 is on the same page as feature vector 21 (m) of the dictionary word 13 is mismatched to feature 22 because it contains a critical point which is not contained in word 12. An examination of tables 1 and 2 indicates that the critical point C_1' is contained in word 13 but not in word 12.

The mismatched critical point is first compared to the correspondence table. If it is not found in the correspondence table then it is stored in a mismatch table. At any latter stage a mismatched critical point is erased from the mismatched table if some word containing the mismatched critical point matches some feature of the dictionary.

In general if word n matches feature m then it is expected that word (n +1) will match feature (m + 1). If feature (m+1) is on the same page as feature m then it is

TABLE 3

REVISED FEATURE VECTOR # 22

CRITICAL POINT	COORDINATE		FEATURE VECTOR	
SEQUENCE #	X-LOCATION	Y-LOCATION	DISTANCE	NORM. ANGLE
C_i	$X_{k,i}$	$X_{k,i}$	$s_{k,i}$	$\alpha_{k,i}$
20	475	350	0.275903	0.3374
21	450	350	0.108797	-.4576
22	410	385	0.360555	0.1203

FEATURE # k=22 REVISED TO D
SIZE = 116.667 ANG MEAN = -0.7827
X-CENTROID=445 Y-CENTROID = 361.667

21	450	350	0.360555	0.1203
22	410	385	0.108797	-.4576
23	385	385	0.275903	0.3374

FEATURE # k=22 REVISED TO D
SIZE = 116.667 ANG MEAN = -0.7827
X-CENTROID=415 Y-CENTROID = 373.333

22	410	385	0.263178	0.4869
23	385	385	0.142178	-1.3906
24	370	345	0.334767	0.9037

FEATURE # k=22 REVISED TO D
SIZE = 96.667 ANG MEAN = .06478
X-CENTROID=388 dY-CENTROID = 371.667

TABLE 4

CRITICAL POINT	COORDINATE		FEATURE VECTOR	
SEQUENCE #	X-LOCATION	Y-LOCATION	DISTANCE	NORM. ANGLE
C_i	$X_{k,i}$	$Y_{k,i}$	$s_{k,i}$	$\alpha_{k,i}$
11	2.7724	11.6666	0.296925	0.3375
12	2.4187	11.3131	0.117094	-.4578
13	1.3580	11.2429	0.388017	0.1203

FEATURE # k = 13 REVISED FEATURE TO P = 13,1
SIZE = 2.6183 ANG. MEAN = 0.07666
X-CENTROID = 2.18303 Y-CENTROID = 11.4075

12	2.4187	11.3131	0.388017	0.1203
13	1.3580	11.2429	0.117094	-.4578
14	1.0043	10.8894	0.296925	0.3375

FEATURE # k = 13 REVISED FEATURE TO P = 13,2
SIZE = 2.6182 ANG MEAN = 0.07666
X-CENTROID = 1.5936 Y-CENTROID = 11.1485

13	1.3580	11.2429	0.291751	1.5340
14	1.0043	10.8894	0.157579	-.3433
15	1.3575	10.1115	0.371073	-1.1907

FEATURE # k = 13 REVISED FEATURE TO P = 13,3
SIZE = 1.74413 ANG MEAN = -.19739
X-CENTROID = 1.23993 Y-CENTROID = 10.7479

14	1.0043	10.8894	0.277421	-.1909
15	1.3575	10.1115	0.083901	0.2752
1	2.6298	8.8382	0.354384	-.0843

FEATURE # k = 13 REVISED FEATURE TO P = 13,3
SIZE = 4.14820 ANG MEAN = -.76957
X-CENTROID = 1.6638 Y-CENTROID = 9.9464

easy to isolate the mismatched point as in the above example. If feature m+1 is not on the same page as feature m or the concept of pages is not used then in order to isolate the mismatched critical point then the feature m+1 and the word n+1 have to be revised into concatened shape vectors with each subshape vector of three measurements. A revised feature vector for feature 22 is shown in Table 3. While the revised words for word 13 are shown Table 4. A comparison of the above two tables again isolates C_1 as the mismatched critical point. It may also be noted from tables that angle being a multivalued function is an unreliable variable for comparison. At the end of this comparison process an intelligent machine would compare other shape parameters like curvature, symmetry etc. It may also try to investigate if the mismatched critical points are due to noise errors, or due to changes in the shape.

CONCLUSION

In this paper, it is shown by example that several of the previously described shape algorithms do not perform well on arbitrary shapes. While each of the methods exhibit specific attributes, it is relatively easy to find shapes that render each of the algorithms useless. This is particularly true when the shapes under comparison are partial shapes.

In particular an example was presented that demonstrated that the Fourier Descriptors method is not suitable as a general method for the recognition of partial shapes. It was also demonstrated that the peaks obtained by the Local descriptors are not independent of rotation.

In the process of defining a shape recognition procedure that would overcome or circumvent the weaknesses of these algorithms, the basic requirements for the shape recognition process were stated. A new concept of treating shapes as vector in shape space was introduced and described. Also two theorems relating to the process of comparing partial shapes to the complete shape were stated and proved.

A new proceedure of determining the critical points of a shape was described. This procedure is named the Adaptive Line of Sight method. In the Adaptive Line of Sight method , the critical point determination is based on a set of coordinate axes that are dependent on the shape being examined. Examples were given that demonstrate that the

procedure produces critical points that are independent of rotation, size, displacment, and correspond closely to those produced by normal human cognitive process. The experimental results indicate that this algorithm , like the human eye, converges to a minimal set of critical points called the segmenting set.

The minimal segmenting set found by this method in rare instances does not coincide with the minimum segmenting set found visually but the critical points turn out to be the same. For instance, a shape that is more than a half-circle, but less than a full circle will lead to ambiguities.

It was demonstrated that measurements between a set of adjacent critical points that were determined by using this method can be used to define feature vectors for shapes. These feature vectors remain the same whether the shape is a partial shape or a more complex whole shape. Moreover, these feature vectors are independent of size, rotation and displacement since they are derived from a set of critical points that are independent of the same quantities. A technique for comparing the feature vectors of a set of shapes is described. The comparison procedure is based on a syntactic technique which will point out whether the shapes are the same, whether one shape is part of a more complex whole shape, or whether the shapes are totally dissimilar.

FLOWCHART FOR THE ADAPTIVE LINE OF SIGHT ALGORITHM

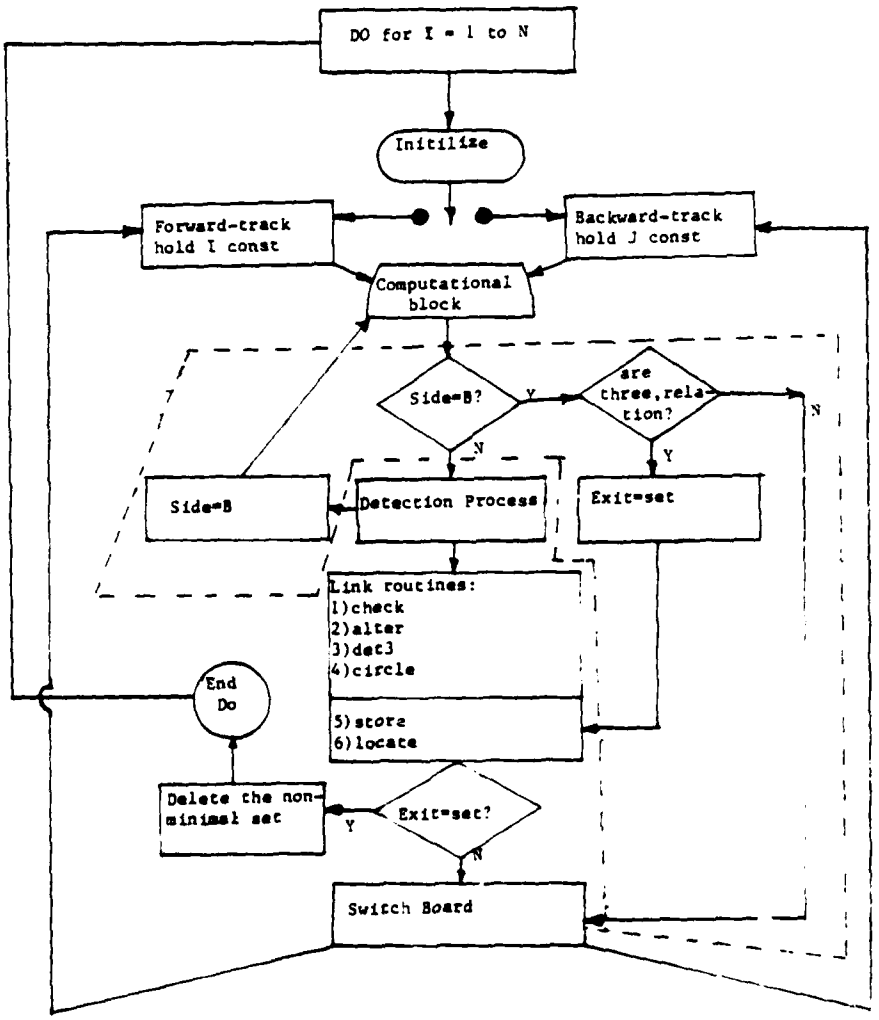


FIG 18. FLOWCHART FOR THE ADAPTIVE LINE OF SIGHT METHOD.

Appendix A

Refering to the flow-chart shown in Fig. 18, Side A refers to the condition when the computations are performed from I to J modulo the number of points in the shape, while side B refers to the condition when computations are performed from J to to I. Forward track or Ftrack denotes to a condition when I is held constant while J is incremented. Backtrack or Btrack denotes a condition when J-1 is held constant while I is decremented.

Initially I and J are always chosen to be adjacent points,first going in the clockwise direction then in the clockwise direction The details of the computational block and the detection block/process are as follows

COMPUTATIONAL BLOCK

Find the equation of the straight line S joining I to J

Find the distance DISTAIJ between points I and J.

Find the equations of the straight lines normal to L and joining every point P in between I and J.

Find the intersection (XINTSEIJ , YINTSEIJ) of each of the normal lines found in the above step with the straight line L

Find the distances DISTXIIJ and DISTYIIJ from point

I to the intersection and from point J to the intersection, respectively for every P.

Find the normal distance NDISTAIJ from every P to the straight line L.

Find the normal vectors from the straight line to every point P.

DETECTION PROCESS

In this block a detect switch is set indicating that a critical point has been detected at I and J-1, if at the first instance, a point is found which is not on the same side of L as other points, or a point cannot be mapped injectively on to the straight line L. The former condition is checked by comparing the magnitude of the sum of every two adjacent normal vectors with the magnitude of those forming the sum ± a THRESHOLD1, while the latter condition is checked by comparing the sum of the distance DISTXIIJ+DISTYIIJ to DISTAIJ ± THRESHOLD2. Where thresholds 1 and 2 are set to account for round-off, truncation, quantization and other errors.

SWITCH BOARD

This is a control block which forces the computations to occur in an alternating sequence FORWARD-TRACK -BACK-TRACK -FORWARD- TRACK.....

REFERENCES

1. Bjorklund, M. C., and Pavlidis, T. "Global Shape Analysis by k-Syntactic Similarity," IEEE Trans., Vol. PAMI-3, No. 2, March 1981, pp. 144-155

2. Davis, L. S., "Understanding Shapes: Angles and Sides," IEEE Trans., Vol. C26, No. 3, March 1977, pp. 236-242.

3. Freeman, H. "Shape Description via Critical points," Pattern Recognition, Vol. 10, pp. 159-166.

4. Fu, K.S. Syntactic Methods in Pattern Recognition. Academic Press, New York 1974.

5. Granlund, G. H., "Fourier Preprocessing for Hand Printed Character Recognition," IEEE Trans., Vol. C-21, Feb. 1972, pp. 195-201.

6. Hall, E. L., and Davis, A., and Casey, M. E., "The Selection of Critical Subsets for Signal, Image, and Scene Matching," IEEE Trans., Vol. PAMI-2, No. 4, July 1980, pp. 313-322.

7. Hermann, R. Differential Geometry and the Calculus of Variations. Academic Press, New York, N.Y. 1968.

8. Johnston, E., and Rosenfeld, A. "Angle Detection on Digital Curves," IEEE Trans., Vol. C-22, 1973, pp. 875-878.

9. Kashyap, R. L., and Chellapa, R. "Stochastic Models for Closed Boundary Analysis, Representation and Construction," IEEE Trans., Vol. IT-27, No-5, Sep. 1981, pp. 627-637.

10. Klingenberg, W. A Course in Differential Geometry. Springer- Verlag, N.Y., 1978

11. Mosimann, J. E. "Size Allometry: Size Shape and Size Variables with Characterization of the Log-Normal and Generlaized Gamma Distribution," J.ASA., Vol. 65, June 1970, pp. 930-945.

12. Nevins, A. J., "Region Extraction from Complex Shapes," IEEE Trans., Vol. PAMI-4, No. 5, Sep. 1982, pp. 500-511.

13. Pavilidis, T. Structral Pattern Recognition. Springer-Verlag, N.Y., 1977.

14. Pavilidis, T. "Algorithm for Shape Analysis of Contours and Waveforms," IEEE Trans., Vol. PAMI-2, No. 4, July 1981, pp. 301-312.

15. Persoon, E., and King-Sun Fu, "Shape Discrimination using Fourier Descriptors," IEEE Trans., Vol. SMC-7, No. 3 March 1977, pp.170-180.

16. Proffit, D. "Normalization of Discrete Planar Objects," Pattern Recognition, Vol. 15 No. 3, pp. 137-143.

17. Richard, C. W., Jr. and Hemami, H. "Identification of Three Dimensional Objects using Fourier Descriptors of the Boundary Curve," IEEE Trans., Vol. SMC-4, July 1974, pp. 371-378.

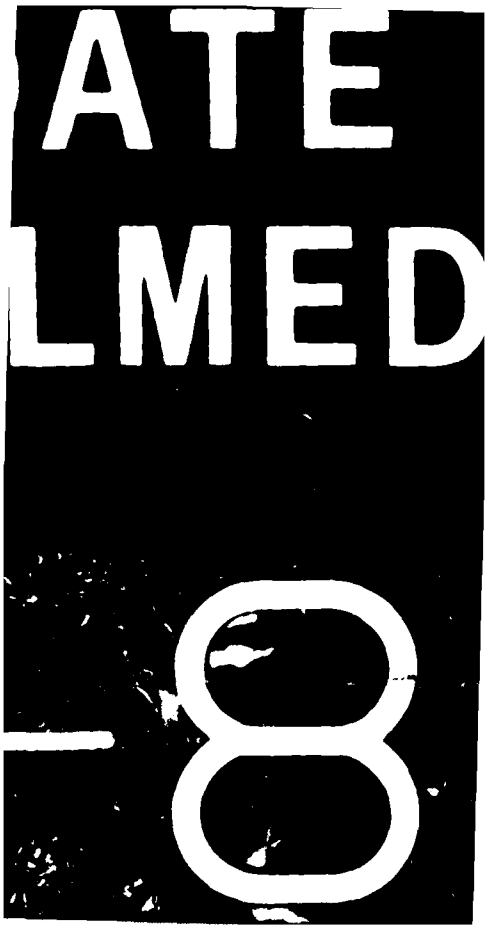
18. Rosenberg, B. "The Analysis of Convex Blobs," Computer Graphics and Image Processing, Vol. 1, 1972, pp. 183-192.

19. Snyder, W. E., and Tang, D. E., "Finding the Extrema of a Region," IEEE Trans., Vol. PAMI-2, No. 3, 1981, pp. 261-269.

20. Thomas, S. J., Louis, R. E., and Malindzak, G.S., "An Algorithm for line Intersection Identification," Pattern Recognition, Vol. 13, No. 2 1981 pp. 159-165.

21. Wallace, T. P., MItchel, O. R., and Fukunaga, K. "Three Dimensional Shape Analysis using Local Shape Descriptors," IEEE Trans., Vol. PAMI-3, No. 3, May 1981, pp. 310-323.

22. Zahn, C. T., and Roskies, R. Z. "Fourier Descriptors for Plane Closed Curves," IEEE Trans., Vol. C-12, March 1972, pp. 269-281.


ATE
LMED
8